情報科学講座 A・2・6



A. 共通基礎理論
A・2 情報論理学

北川敏男編

確率論的エントロピー

―情報理論の函数解析的基礎 1―

梅垣壽春
大矢雅則 共著

共立出版株式会社

“情報科学講座”の序

情報科学は機械・生体および人間社会における情報の生成・伝達・改造・蓄積・利用についての一般原理を攻究する基幹科学である．情報理論の建設，情報現象の解明，情報方式の開発の三つの領域にわたり，相互間の緊密な協力によってその発展をはかることは，現代科学技術の進歩にとって，必須の要請となっている．

21世紀の人類のビジョンは，情報革命のもとに築かれなければならない．20世紀前半における物理科学の進展は，やがて生物科学，人文科学，社会科学に大きな影響を及ぼし，これらの科学分野の飛躍的発展が期待されている．この時代において，それらの相互間の連結は，情報科学を通じて行なわれるべきものであり，情報科学の進歩は，これらの諸分野の発展に強力な推進力となるものと期待される．科学・技術の研究が，人類社会に占める役割は年とともに加重しているが，科学・技術の研究のために共通の基盤を提供するものは，計算機といい，ドキュメンテーションといい，情報科学の負担すべき任務に属する．

情報科学の組織的研究体制を整備するとともに，情報科学について系統的な学習を行ない，広範な教養を培い，わが国における情報科学の水準を世界のそれに遅れないようにすることは，今日の急務である．

このような趣旨から，情報科学講座（刊行当初全 63 巻）を刊行しようというのである．講座の意図するところは，情報科学の体系的集成であるから，理論・素子・組織・生体情報・装置の全分野にわたり，基礎的な解説から，第一線の研究の紹介にまで及ぶように努めた次第である．

わが国の情報科学の水準が世界をリードし卓越する日の来ることを待望しつつ，この目的のために，情報科学講座がいささかなりとも寄与できることを，心から念願するものである．

(1966 年 9 月記，1977 年 1 月一部変更)

編集委員一同

序

エントロピーは，現代，いろいろな分野に通底して，広く採用される基本概念の一つになっている．それは，1980 年代の現在では，Georgescu–Roegen によって経済過程論に，Henderson によって未来論に，Attali によって「言葉と道具」の考察に，Rifkin によって経済学から世界観の確立に，適用されている．エントロピーの概念を正確に理解し，適切に利用すること，さらにいうならば，来るべき時代の要請に対処すべくこの概念の拡充をはかることが必要と思われる．エネルギーとエントロピーという基本対概念に対して，さらに何らかの基本概念を追加すべきではないか．このような予感さえ抱かしめるような時代になっている．

このような時点において，エントロピー概念への数学的接近を近代的な函数解析学の立場から解説した本書を，情報科学講座 A·2·6 として公刊できることになったことは，まことに時宜をえたものかと思われる．

本書の著者の一人，東京工業大学理学部梅垣壽春教授は，多年にわたり情報理論およびエントロピー論に多くの輝かしい業績をあげられたエキスパートである．今回，東京理科大学助教授大矢雅則博士のご協力のもとに，2 巻にわけて，この方面の最近の成果までにわたって懇切丁寧に解説していただき，確率論的および量子論的エントロピー論の全貌を明らかにしていただくことになったのは編集者として深謝に堪えない．本巻で援用される函数解析学などの数学的基礎については，特に付章を設けて，必要最小限の範囲で簡潔な解説が準備され，読者のための便がはかられている．この準備的知識を身につけるのは，相当の負担を一般の読者に要請するものではあろうけれども，その大要を捉ええたうえで，その都度付章を参照しつつ進まれたい．これによって本書全体の解説が，透徹した論旨のもとに，極めて明快簡潔に記述され展開されているの

を，読者は了解されるであろうし，負担加重の労は充分報いられるものと確信する次第である．

エントロピーの概念は，周知のように物理学の発展の歴史のなかに生誕されたものである．19 世紀の中葉，熱機関の効率問題に本格的に取り組んだ物理学者および工学者の努力は，Carnot 原理，Lord Kelvin の熱理論をうみ，Clausius の熱力学の二つの基本法則として結実すると共に，Boltzmann を経て Gibbs の統計力学の樹立を導いた．ここに本書に述べられているような Clausius エントロピー，Boltzmann エントロピーが出現し，エネルギーの概念と共に，エントロピーの概念の生誕がまず物理学の世界で記録されたのである．

それから約 1 世紀を経て，今世紀の中葉以来，エントロピー劇の舞台は情報科学へ移る．Wiener は，物理学における Gibbs 統計力学と数学における Lebesgue 積分との両刀を武器として，通信と制御の理論としてのサイバネティックスをつくり，秩序に対する混沌（chaos）の確率過程論を展開した．他方，通信理論への統計学的接近として，Shannon は美麗な情報理論を創設したが，ここに確率論的 Shannon のエントロピーの概念が主役を演ずる．Wiener の混沌と Shannon のエントロピーとはいわば双生児である．これらを統合した通信（Communication）と制御（Control）の理論のほかに，Turing および von Neumann の計算（Computation）に関する理論が樹立されて，3 C 技術の基礎ができ，三者の綜合としての情報処理技術の確立・普及・発達と共に，情報科学が 1960 年代より開花する状況となったのである．

Shannon の通信模型，エントロピーの概念は，直ちに Osgood らにより，コミュニケーション科学の基礎として広く援用されている．また不確実事象に対応する場面においてエントロピー概念が基本的な役割を果すことは，国沢清典教授らによって，詳細に究明されたところである．さらにエネルギー問題，環境問題が全世界的な規模において，切実深刻な政治・経済的な課題となるにつれ，前述のように諸家によりとりあげられ，「エントロピーの法則」が，往々にして拡大解釈のもとではあるが，援用されている状況にさえなってきた．まことに，何らかの体系的理論を待望する時代の要請は切実である．

他方顧みれば，情報科学の論理の要請する数学型式はまだ充分には開発されていない現状である．また生体情報学に視点を移せば，Boltzmann により指摘され，のち Schrödinger によって提唱された「負のエントロピー」摂取問題をふくめて，エントロピー概念の再検討も要請されている．文化理論の記号的接近における秩序と混沌，中心と周縁，ここにもエントロピー論のアナロジーというよりも，より深い省察の機縁が提示されている．

このような世界情勢と学界風景とのもとにおいてであるが，読者は，心静かに本書によって明澄な数学理論を学習する時間をもち，他日の飛躍の基礎づくりにそなえていただきたいと思う．この方面の数学理論は，上述のように近代解析学の精華を利用して構築されるものである．他方，例えば Hilbert の第 13 問題解決への ε-エントロピー理論の応用は Kolmogorov によって本書の定理 4·6 に見られるように，新しい数学的接近法開発の端緒をも示唆している．はじめアメリカに産れ，やがて強じんな確率論の伝統的基盤をもつソ連において，エントロピー理論は厳密な再構成のもとに多彩な展開を示したことは，本書に如実に示されている．しかしひとりソ連のみでなく，わが国をふくめて，広く国際的な共鳴と関心とのもとに発展したのがこの方面の研究である．豊富な引用文献を通じてこれらの多くの研究への接近するパスが本書によってつくられたのは学界のため幸甚なことである．多くの分野の好学の士の精読を期待してやまない．

1983 年 10 月

編集者　北　川　敏　男

自　　序

C. E. Shannon は 1948 年に論文「通信の数学的理論」を発表し，通信理論に飛躍的進歩をもたらしたが，同時にそれが，今日，情報理論と名づけられ，いわゆる現在の情報化社会における「情報」という概念の根幹を与えたのである．

現在，情報理論が対象とする範囲は純粋から応用にわたる数学分野・量子統計力学・通信・エレクトロニクス関係はもとより，言語理論・心理学・生物科学・経済学・ゲーム論等，さらに社会科学などにも及んでいる．情報理論において「情報」の量的把握としてまず「エントロピー」が登場する．このエントロピーなるものが，曖昧さ，予測などの不確かさ，また，秩序の度合いなどを計る合理的スカラー量としての有効性をもち，さらにビットの単位にもなっており，総合して上述の諸分野を推進する源となっている．それ故に情報理論が現在各方面から研究の対象として注目を浴びるようになって来たのである．

情報理論に関する書物は多数出版されているが，いずれも類似な工学的色彩が強いように見受けられる．本書では，これを数学的色彩を強くし，数学の発展領域として情報理論を捉えようという，そのような計画の下に記述されている．

学問とは混沌に目鼻立ちをつけることが，その主たる任務であるといわれるが，「エントロピー」こそ，混沌そのものを定量化する最初でかつ目下のところ，最も有力な概念である．対象となる情報源という一つの「系」——何らかの重みをもつものの集団，それが確率事象系である場合もあり，量子力学系であることもある——を記述する状態の，無秩序の度合い，すなわちその不確定さの大きさを表すものがエンロピーというスカラー量なのである．この無秩序を顕すということは裏を返せば，系の秩序の程度を顕していると考えてよい．

一般に，不確定さの大きな系からある情報を得たならば，その情報は不確定さの小さな系から得た情報よりも価値があると考えられる．Shannon は基本的に，この考えの上に立って不確定さの大きな系はより多くの情報を含んでいる，したがってエントロピーはその系のもつ情報の量を表していると見なし，情報理論の礎を築いたのである．

さて，エントロピーという概念は 1865 年 Clausius によって導入され，Boltzmann, Gibbs, Nyquist, von Neumann, Wiener, Shannon, Kolmogorov 等々によって 20 世紀の中頃までに種々な方法によって研究されてきたが，現在では，情報科学においてはもとより，数学，物理学などにおける重要な不変量であり，基本的な概念として必要欠くべからざるものとなってきた．また，前述したようにエントロピーは種々な生体の振舞い，社会現象などの解析にも大きな威力を発揮している．本書ではこのエントロピーの数学的基礎づけがその主たる目的である．

本書を理解するためには数学の基礎知識をある程度必要とする．そのために付章１～５を設けてある．この付章は大学の数物系，電気系，制御システム科学系などで必要とする数学の予備知識——およそ学部２，３年前期まで——が網羅されているので，読者の所有する数学知識に応じて取捨選択されたい．例えば，情報関連学科では情報数学というような呼称で設けられている授業はこれに該当しており，数学科や応用数学科では集合・位相・代数・測度・Hilbert 空間入門などがこれにあたる．

本文では，第１章はエントロピーという概念の発生に遡って Clausius エントロピーから始める．この章の中心は Shannon エントロピーである．第２章は Kolmogorov 理論である．ここまでは付章の知識はほとんど不要である．エントロピーの詳しい内容はさておいて，いわゆる情報理論のみを目的とするならば，第１章，§3，4から始めて第２章を読み，いきなり第５章に進み，第６章の §1，2 という風に進めばよい．しかし，本書の目標とするところは，さらにエントロピーの解析にあるので，他の部分も当然読まれることが望ましい．すなわち，第１章の§1の基本的定理，第４章の ε-，位相，fuzzy-エント

ロピー．また第6章では Fourier 解析を用いる信号解析を論じ，Heisenberg の不確定性原理が Hilbert 空間の Schwarz 不等式の直接の応用として導かれることや，エントロピーとの関連も加えて説明してある．

本書は当初一冊として計画されていたが，情報理論を数学的な見地に立って展開するため付章も増え，またこの数年来発展して来た光通信理論などの関連で，量子情報理論の基礎となる量子論的エントロピー，同チャンネル，非可換作用素系のエントロピー，非可換確率論なども論じる必要性が生じ分量が増大し，応じて章の数も増えた．このため2分冊に分け，主として確率事象系におけるエントロピーを中心とする部分（第1章～第5章）に第6章信号解析を加え，それに付章1～5を付けて第1巻とした．これが本書の構成である．また，数学的には非可換解析，具体的には量子力学的エントロピーの理論を中心として，第7章から第 12 章および付章1～4（第1巻とは別）からなるものを第2巻「量子論的エントロピー」とした．これは本書に続いて出版される予定である．

本書は直前に出版された国沢清典先生の情報理論と内容的には共通点もあり，本書に詳しくは述べなかったエントロピーの具体例やその最大化の議論も同書に詳しいので参照されればなお一層望ましい．

北川敏男先生からこの講座の基礎的な部分として情報理論を数学的見地に立って書かないかとお推めを受けてすでに久しい．その後，右顧左眄して長年月が経過し今日に至ってしまった．この点，北川先生には深くお詫び申し上げねばならない．この歳月の間，情報理論も進展し，その中心的概念であるエントロピーも様々な手法によって研究され，このところ，エントロピーのブームが出現した．著者らは数学的に厳正な見地に立って，このエントロピーに関する主要な数理構成を本書とそれに続く第2巻によって汎く記述したつもりである．

また共立出版社の佐藤邦久氏には，本書が計画されてより今日に至るまで，文字通り永い年月にわたって著者の研究室を訪問され，内容についての詳しい打ち合わせを重ねられたが，同氏の熱意に心から敬意を払うものである．予定

の分量をはるかに超過したにもかかわらず筆者の希望を実現していただいた．本書を世に出しえたのも同氏のご尽力のたまものである．本書は広汎にわたる概念を駆使するため多種多様の活字を用いざるをえない．この困難は編集担当の豊田秀一氏の努力によってすべて解決できた．このことを重ねてお礼申し上げたい．また原稿の浄書に関して東京工業大学理学部および東京理科大学理工学部両大学の情報科学専攻大学院生８名の諸君の協力を得たことも付記して謝意を表したい．

昭和 58 年　盛夏

梅　垣　壽　春
大　矢　雅　則

目　　次

第5章 通信における数理解析

第6章 信号解析

付章 数学的基礎

凡　例（本書を読むために）

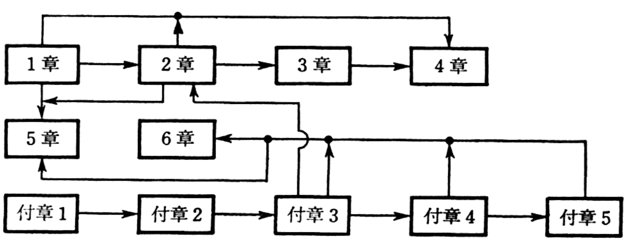
（1）　記号“≡”は左を右で定めることを意味し，“$\overset{\text{def}}{\Longleftrightarrow}$”は左を右で定義することを意味する．

（2）　“A⟺B”はAとBは同値（同等）であると読み，“A⟹B”はAならばBであると読む．

（3）　文中では付章を引用する場合は，例えば，（付章 1.1）あるいは（付章 1.1 参照）などと記す．

（4）　参考文献の引用は，例えば Shannon の［1］は，［S. 1］としてある．

（5）　各章の相互関連は次のようになっている．



第2巻 '量子論的エントロピー' の目次

第1章　熱力学的エントロピーと情報論的エントロピー

この章では，Clausius によって導入された熱力学的エントロピー，および Boltzmann によって導入された統計力学的エントロピーの簡単な解説を行い，これらのエントロピーがどのように係わっているかを見ておく．さらに，離散的事象系のエントロピー，いわゆる Shannon エントロピーを導入し，その主要な性質に言及し，それと熱力学的エントロピーとの関連について議論する．

1.1　Clausius エントロピー

ある物理系が外界と接し，エネルギーのやり取りを行っているとする．エネルギーは状態量（すなわち，その変化はとる道によっていない）とし，系の内部エネルギーの変化を dE で表し，系が吸収した熱量を $d'Q$，系が外界からされた仕事を $d'W$（たとえば，外界の圧力を p，系の体積変化を dV とすると，$d'W=-pdV$）で表すと，

$$dE=d'Q+d'W$$

の関係がある．これが**熱力学第一法則**である．なお，ここで $d'Q, d'W$ は一般に全微分でない（つまり，Q, W は状態量とは限らない）ので，ダッシュを付けてその増分を表している．

系が一様で，その部分部分で圧力 p，温度 T が異なる値をとることがない場合は（そうでないと巨視的な観測は困難になる），系の状態を表す変数 p, T, V の間にはある関係式 $F(p, T, V)=0$ が成立する．系の巨視的な性質を示すこの式は，しばしば**状態方程式**とよばれている．たとえば，n モルの理想気体は，よく知られているように，$pV=nRT$（R はある定数）という関係式を満たしている．

さて，このように系全体が一様であるとき，巨視的に見て静止している，い

わゆる熱平衡状態において系をゆっくり変える（ある意味で思考上の）変化を**準静的変化**というが，この変化は次の意味で可逆変化（過程）である．すなわち，系をある状態 a から他の状態 b へ変化させ，その後 b から a へもどすといった変化を外界へ何の影響も残さず行える場合，変化 $a\to b$ を可逆変化というのである．

ところで Clausius は，**熱力学第二法則**（たとえば，「高温の物体から低温の物体へ熱を移す過程は不可逆である」（**Clausius の原理**））から，状態がもとにもどる1サイクルの間に，熱源の温度を T とすると，

$$\oint\frac{d'Q}{T}\leq 0$$

という不等式が成立することを示した．ここで等号は可逆変化のとき成立するが，このとき $d'Q/T$ は全微分となるので，これを dS とかき，S を **Clausius**（**熱力学的**）**エントロピー**とよぶ．

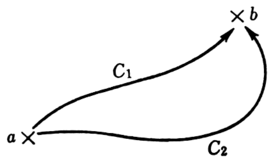


図 1.1

今，ある系が状態 a から他の状態 b に可逆的に移る場合（C_1）と，不可逆的に移る場合（C_2）を考える．

このとき，C_1 の逆向きの道 $-C_1$ と C_2 を合わせると，ひとつの不可逆サイクルができるから，Clausius の不等式より

$$\int_{-C_1}\frac{d'Q}{T}+\int_{C_2}\frac{d'Q}{T}<0.$$

ところで，可逆変化の部分（$-C_1$）では

$$\int_{-C_1}\frac{d'Q}{T}+\int_{C_1}\frac{d'Q}{T}=0$$

であり，この部分では T は系の温度でもあるから，

$$S(b)-S(a)=\int_a^b dS\ \ >\int_{C_2}\frac{d'Q}{T}$$

となり，不可逆過程においては，熱源のエントロピー増分より系のエントロピー増分の方が大きいことがわかる．さらに，系が外から熱を受けない（断熱）不可逆過程の場合は，$d'Q=0$ であるから，

$$S(b)-S(a)>0$$

となり，外界からのエントロピーの出入りはないが，系の内部でエントロピーは増大していることがわかる（エントロピー生成）.

このことと Clausius の不等式が熱力学第二法則から導かれているという事実から，エントロピーが増加する方向が，自然現象の変化の方向，いわゆる時間の向きを示していると考えることもできるのである.

1.2 Boltzmann エントロピー

あらゆる物理系は微視的な粒子が多数集まってつくられていると考えられるから，系の巨視的な熱力学的性質を粒子の微視的な運動論的立場から説明することが望ましい.

この問題を最初に提起したのが Boltzmann である. Boltzmann は多数の粒子からなる巨視的な系では，各粒子のもつエネルギーの差より，外界からの系に対する"ゆらぎ"（攪乱のエネルギー）の方が大きいと見なしうるので，完全な力学的記述はそれほど意味がなく，むしろ系の統計的な取り扱いの方がより重要であると考えた.

今，対象としている物理系が m 個の部分系に分けられ，各部分系はある物理量 E_k $(k=1,2,\cdots,m)$（たとえば，エネルギー）によって特徴付けられているとする.

このとき，系全体の粒子数を N，k 番目の部分系の粒子数を N_k とすると，外界からの粒子の出入りがなければ，

(1) $N=\sum_{k=1}^{m} N_k$

が成立し，さらに k-系の状態数（エネルギー E_k をもつ状態の数（縮退度）など）を g_k とすると，系全体のとりうる状態の総数 W は，

(2) $W=\dfrac{N!}{N_1!\cdots\cdots\cdots N_m!}g_1^{N_1}\cdots\cdots\cdots g_m^{N_m}$

で与えられる.

ところで，この W が最大となる分配 $\{N_1,\cdots,N_m\}$ において系は最も安定（熱平衡）であると考えられるが，以下この分配 $\{N_1,\cdots,N_m\}$ を次の付加条件

(i),(ii) のもとで求めてみよう：

(i)　$N=\sum_{k=1}^{m} N_k=$一定,

(ii)　$E=\sum_{k=1}^{m} E_k N_k=$一定.

まず，よく知られた Stirling の公式を証明しておこう.

〈補題 1.1〉 Stirling の公式：ある $0\leq\theta_N<1$ に対して,

$$N!=\sqrt{2\pi N}\ N^N \exp(-N+\theta_N/12N).$$

証明　$a_N=N!/\sqrt{2\pi N}\ N^N e^{-N+1/12N}$,　$b_N=N!/\sqrt{2\pi N}\ N^N e^{-N}$ とおくと, $a_{N+1}>a_N$, $b_{N+1}<b_N$ がわかり，さらに, $b_N/a_N=e^{1/12N}\to 1$ より，$c=\lim a_N=\lim b_N$ が存在する．$a_N<c<b_N$ であるから，ある数 $0\leq\theta_N\leq 1$ を使って $c=e^{-\theta_N/12N}b_N$ とかける．ここで,

$$\textbf{Wallis の公式}\quad \lim_{N\to\infty} 1/2N(N!2^{2N}/(2N)!)^2=\pi/2$$

を使うと，$c=1$ が出る．したがって，$e^{-\theta_N/12N}b_N=1$. ゆえに,

$$N!=\sqrt{2\pi N}\ N^N \exp(-N+\theta_N/12N).\ \blacksquare$$

粒子の数 N, N_k は通常非常に大きな数であるから（これを記号 $N\gg 1, N_k\gg 1$ で表す)，$\log N!=N\log N-N$,　$\log N_k!=N_k\log N_k-N_k$ と考えてよい．この近似式を **Stirling の近似式**という.

Stirling の近似式が成立する場合に，W を最大とする分配 $\{N_k\}$ を求めよう．W が最大であることと，$\log W$ が最大であることとは同値であるから，条件(i),(ii) を満たすように Lagrange の未定係数法を適用すると,

$$\delta(\log W+\alpha\sum_k N_k-\beta\sum_k E_k N_k)=0.$$

ここで，α, β は Lagrange の未定係数である．上の変分より，各 k に対して,

$$\alpha-\beta E_k+\frac{\partial}{\partial N_k}\log W=0$$

が得られる．ここで，$W=(N!/N_1!\cdots N_m!)g_1{}^{N_1}\cdots g_m{}^{N_m}$ であることから，$N_k\gg 1$ のときは,

$$\frac{\partial}{\partial N_k}\log W=\log g_k-\log N_k$$

となり，したがって，$N_k=g_k\exp(\alpha-\beta E_k)$.

ここで，$Z=\sum_k g_k\exp(-\beta E_k)$ とおくと，条件（i）$N=\sum_k N_k=e^\alpha Z$ より α が求まり，さらに E_k, E, g_k を知ることができれば，条件（ii）より β も求まる．実際，粒子の運動論的考察により，**Boltzmann 定数** k（$=1.38\times10^{-16}$ erg/°K）と系の温度 T を使って，$\beta=1/kT$ と表されることがわかっている．したがって，$N_k\gg1$ の場合，W を最大にする分配 $\{N_j\}$ は

$$N_j=\frac{g_jNe^{-E_j/kT}}{Z}$$

で与えられる．

この $\{N_j\}$ において，$\log W$ を計算すると，

$$\begin{aligned}\log W&=N\log N-\sum_k N_k\log N_k+\sum_k N_k\log g_k\\&=N\log N-\frac{N}{Z}\sum_k g_ke^{-\beta E_k}\log\left(\frac{N}{Z}e^{-\beta E_k}\right)\\&=N\log Z+\frac{N}{Z}\beta\sum_k g_kE_ke^{-\beta E_k}\\&=N\log Z+\beta E.\end{aligned}$$

したがって，

$$\begin{aligned}\frac{d}{dT}(k\log W)&=\frac{kN}{Z}\frac{dZ}{dT}-\frac{E}{T^2}+\frac{1}{T}\frac{dE}{dT}\\&=\frac{1}{T}\frac{dE}{dT}.\end{aligned}$$

系の体積 V は一定であると見なせるから，定積比熱を c_V とすると，$dE/dT=c_V$ であり，また，$d'Q=c_VdT$ であることから，

$$d(k\log W)=\frac{d'Q}{T}=dS$$

となり，

$$S=k\log W$$

が導き出された．こうして熱力学的エントロピー S は，Boltzmann の熱力学

的重率 W によって定めることができたのである．$S=k\log W$ は通常 **Boltzmann のエントロピー**とよばれている（このよび名は Pauli による）．

1.3　完全事象系と Shannon エントロピー

n 個の元からなる集合 $X=\{x_1, x_2, \cdots, x_n\}$ $(n\leq\infty)$ と，その各元の生起する確率 $p=\{p_1, p_2, \cdots, p_n\}$ (i.e.，$\sum_{k=1}^{n} p_k=1$，$p_k\geq 0$) の組 (X, p) を**完全事象系**という．これは物理的な観測という観点からすれば，状態 p において，事象 $x_k\in X$ を観測する確率が p_k であるということである $(p_k=Pr(x_k))$．

ここでトランプ 52 枚の中から1枚を抜くという操作を考えてみよう．この場合，あるカード（ひとつの事象）が出る確率を問うと，それは 1/52 である．ところで，52 本のくじの中に当たりくじが1本入っていて，それを引く場合を考えてみると，くじが当たる確率はこの場合もトランプの場合と同様に 1/52 である（当たらない事象の確率は 51/52）．しかし，生起する事象をこの2つの例において見てみると，トランプの場合は 52 個，カードの種類によって異なる事象があり，一方，くじの方は当たるか当たらないかの2個の事象のみである．したがって，くじ引きの場合は，引くという行為に対する結果の予測がほとんど確実にできる．すなわち，事象の予知に関する**不確定さ**は，トランプの場合の方がはるかに大きいということになる．不確定さが大きいということは，その大きい系からある情報を得たならば，その情報は不確定さの小さな系から得るものよりも意味があるということである．では，こういった“不確定さ”はどのように数量化しうるであろうか？

今，ある事象が等確率で起こる2つの完全事象系 $(X, p)=(\{x_k\}_{k=1}^{n}, \{1/n\})$，$(Y, q)=(\{y_k\}_{k=1}^{m}, \{1/m\})$ を考えてみよう．この2つの事象系が独立であるとき，それらの**複合事象系**は，

$$(X\times Y, p\times q)=(\{(x_i, y_j)\}_{ij},\ \ \{1/nm\})$$

と表すことができる．このとき複合事象系 $X\times Y$ のもつ不確定さは，X と Y とが独立であることから，X の不確定さと Y の不確定さの和であるとするのは自然であろう．事象系 X の不確定さをその場合の数の函数として $f(n)$ で表

すと，Y のそれは $f(m)$ となり，先の議論から $X\times Y$ の不確定さは，

（1） $f(nm)=f(n)+f(m)$

となる．さらに，生起する事象がただひとつしかないときの不確定さは，当然0のはずである．つまり，

（2） $f(1)=0.$

さらに，$n\geq m$ のとき

（3） $f(n)\geq f(m)$

が成立することが必要である．以上3つの条件を満たす函数 f は

（4） $f(n)=\lambda\log n$

と定まる．ここで λ は log の底を決める定数と考えてよい．$\lambda=\log_2 e$ ととると，$f(n)=\log_2 n$ となり，とくにこのとき，不確定さは $f(n)$ **ビット**であるという．しかし，以下我々は，とくに断らない限り $\lambda=1$ ととることにする．

（1）～（3）を満たす $f(n)$ が（4）の形に表せることを証明しておこう．

r 個の等確率事象からなる m 個のたがいに独立な完全事象系を $X_1,\cdots,X_m$ とすると，複合事象系 $\prod_{k=1}^{m}X_k$ に対する不確定さは，性質（1）より $mf(r)$ である．複合事象系の事象の数は，r^m 個であるから，

$$f(r^m)=mf(r) \tag{1.1}$$

となる．同様に他の自然数 s,n に対して，

$$f(s^n)=nf(s).$$

ここで，$r>1$，$s>1$，n を任意に与えられた自然数とし，m を不等式

$$r^m\leqq s^n<r^{m+1} \tag{1.2}$$

によって定まる整数とすると

$$m\log r\leqq n\log s<(m+1)\log r.$$

したがって，

$$\frac{m}{n}\leqq\frac{\log s}{\log r}<\frac{m}{n}+\frac{1}{n}. \tag{1.3}$$

性質（3）より

$$f(r^m)\leqq f(s^n)\leqq f(r^{m+1}),$$

ゆえに，

$$mf(r)\leqq nf(s)\leqq(m+1)f(r).$$

これから，

$$\frac{m}{n} \leqq \frac{f(s)}{f(r)} \leqq \frac{m}{n} + \frac{1}{n}. \tag{1.4}$$

(1.3) と (1.4) より，

$$\left| \frac{f(s)}{f(r)} - \frac{\log s}{\log r} \right| \leqq \frac{1}{n}, \qquad \forall n \in \boldsymbol{N}.$$

したがって，

$$\frac{f(s)}{\log s} = \frac{f(r)}{\log r}.$$

r, s は任意であったから，すべての n ($\geqq 2$) に対して

$$f(n) = \lambda \log n$$

とかける．

さらに，一般の完全事象系 $p = \{p_k\}_{k=1}^{n}$ に対する不確定さ $S(p)$ は，前述の条件（1）〜（3）と類似の条件のもとで（詳しくは〈定理 1.2〉参照），

$$S(p) = -\sum p_k \log p_k \quad (\text{ただし，} 0 \log 0 = 0) \tag{1.5}$$

で与えられる．この量は系 (X, p) のもつ**不確定さ**，**乱雑さ**，**あいまいさ**などを表す量であり，(Shannon の) **エントロピー**とよばれている．これは $S(X)$ とかき表されるときもある．

「**p の不確定さの解消 $\Longleftrightarrow$ p に含まれる情報の獲得**」

と考えられるので，量 $S(p)$ は系の有する情報量を表しているものと見なすことができる．

ここで，以下の議論のためにいくつかの記号および概念を導入しておこう．

（1）　事象の集合 X における要素 $x \in X$ の生起確率は，しばしば $p(x)$ とかくことがある．とくに，複合事象系 $X \times Y$ の要素 (x, y) の生起確率は $p(x, y)$ とかく．

（2）　n 個の事象のとる状態（確率分布）$p = \{p_1, \cdots, p_n\}$ の集合を Δ_n で表す．つまり，$\Delta_n \equiv \left\{ p = \{p_k\}_{k=1}^{n};\ \sum_{k=1}^{n} p_k = 1,\ p_k \geqq 0 \right\}$ である．Shannon エントロピー $S(p)$ は，各事象の生起確率を明記する必要があるときは，$S(p) = S(p_1, \cdots, p_n)$ とかくこともある．

（3） 複合事象系の状態 $\{p_{ij}=p(x_i, y_j); x_i\in X, y_j\in Y\}$ のエントロピーは，$S(X, Y)=-\sum p_{ij}\log p_{ij}$ で与えられる．また，$y\in Y$ が起こったという条件のもとで $x\in X$ が起こる**条件付確率** $p(x|y)=p(x, y)/p(y)$ を使って，**条件付エントロピー**が

$$S(X|Y)=-\sum_{y\in Y}\sum_{x\in X}p(y)p(x|y)\log p(x|y)\geqq 0 \tag{1.6}$$

と定められる．

この条件付エントロピーは，X と Y がたがいに独立なとき，$p(x, y)=p(x)p(y)$ であることより $S(X)$ に等しくなる．一般に，条件付エントロピー $S(X|Y)$ は，Y を知った（観測した）後で，まだ X に残っている不確定さを表す量である．

（4） Y を観測することによって得られる X の情報量は**相互エントロピー**（**相互情報量**）といわれ，

$$I(X, Y)=S(X)-S(X|Y) \tag{1.7}$$

で与えられている．

（5） ある完全事象系 X の 2 つの状態 $p, q\in\varDelta_n$ において，p の q に対する相対的不確定さともいうべき量を**相対エントロピー**といい，

$$S(p|q)=\sum_{k=1}^{n}p_k(\log p_k-\log q_k) \tag{1.8}$$

で表す．

〈定理 1.2〉 各エントロピーは次の性質を有する．

（1） 任意の $p, q\in\varDelta_n$ に対して，$S(p|q)\geqq 0$.

（2） 任意の $p\in\varDelta_n$ と重確率行列 $A=(a_{ij})_{i, j=1,\cdots,n}$ (i.e., $\sum_i a_{ij}=\sum_j a_{ij}=1$, $a_{ij}\geqq 0$) に対して $q=Ap$ とおくと，$S(q)\geqq S(p)$.

（3） $S(X, Y)=S(X)+S(Y|X)=S(Y)+S(X|Y)$.

（4） $S(X, Y)\leqq S(X)+S(Y)$.

（5） $S(X|Y)\leqq S(X)$.

証明 （1）：任意の $x>0$ に対して，$x-1\geqq\log x$ であるから，

$$S(p|q)=\sum_k p_k \log\frac{p_k}{q_k} \geqq \sum_k p_k\left(1-\frac{q_k}{p_k}\right)=0.$$

（2）：函数 $\eta(x)=-x\log x$ は $x>0$ で凹であるから，

$$S(q)=\sum_i \eta(q_i)=\sum_i \eta\,(\sum_j a_{ij}p_j) \geqq \sum_i\sum_j a_{ij}\eta(p_j)=\sum_j \eta(p_j)=S(p).$$

（3）：$S(X, Y)=-\sum_{x\in X, y\in Y} p(x, y)\log p(x, y)$

$$=-\sum p(x, y)\log p(x)p(y|x)$$

$$=-\sum p(x, y)\log p(x)-\sum p(x, y)\log p(y|x)$$

$$=S(X)+S(Y|X).$$

同様にして，$S(X, Y)=S(Y)+S(X|Y)$ が出る．

（4）：$S(X)+S(Y)=-\sum_{x\in X} p(x)\log p(x)-\sum_{y\in Y} p(y)\log p(y)$

$$=-\sum_{x\in X, y\in Y} p(x, y)\log p(x)p(y)$$

$$\geqq -\sum p(x, y)\log p(x, y)=S(X, Y).$$

（5）：（3），（4）より明らか．▌

註　上記の不等式の等号はそれぞれ以下の場合において成立する（証明は読者自ら試みよ）．

（1）　$p=q$ (i. e., $p_k=q_k$, $\forall k$).

（2）　$q_k=p_{\pi(k)}$（π は $\{1, 2, \cdots, n\}$ のある置換）．

（4）　と（5）X と Y とが独立（$X \perp Y$ とかく）．

1.4　Shannon エントロピーの特徴付け

エントロピー $S(p)=-\sum_{k=1}^{n} p_k \log p_k$ はどのような特徴をもっているか，またそれはどのように特徴付けられるかを調べよう．

(S. K.)　Shannon-Khinchin の公理系

(S. K. 1)　函数 $S:\bigcup_{n\geq 2}\varDelta_n\to \boldsymbol{R}^+$ は連続であり，$S(1/n, 1/n, \cdots, 1/n)=\max\{S(p)\,;\,p\in\varDelta_n\}>0$.

(S. K. 2)　$S(p_1, \cdots, p_n, 0)=S(p_1, \cdots, p_n)$.

(S. K. 3)　$p_k=\sum_{j=1}^{m_k} q_{kj}$（$m_k\geqq 2$,　$k=1, 2, \cdots, n$）とかけるとき，

$$S(q_{11}, \cdots, q_{1m_1}, \cdots, q_{n1}, \cdots, q_{nm_n})$$
$$= S(p_1, \cdots, p_n) + \sum_{k=1}^{n} p_k S(q_{k1}/p_k, \cdots, q_{km_k}/p_k).$$

(F.) **Faddeev の公理系**

(F.1) 函数 $f(p)=S(p, 1-p)$ は区間 $[0, 1]$ 上で連続で，少なくとも一点 $p_0 \in [0, 1]$ で $f(p_0)>0$.

(F.2) $\{p_k\} \in \Delta_n$ の任意の置換 $\{p_k'\} \in \Delta_n$ に対して，

$$S(p_1, \cdots, p_n) = S(p_1', \cdots, p_n').$$

(F.3) $p_n = q + r$, $q \geqq 0$, $r>0$ とするとき

$$S(p_1, \cdots, p_{n-1}, q, r) = S(p_1, \cdots, p_n) + p_n S(q/p_n, r/p_n).$$

〈**定理 1.3**〉 (S.K.) $\Longleftrightarrow$ (F.) $\Longleftrightarrow$ $S(p) = -\lambda \sum_k p_k \log p_k$.

この定理によって，函数 $S(p)$ が特徴付けられるのである.

証明 ［第1ステップ］: (S.K.) ⇒ (F.) を証明する.

(S.K.1) ⇒ (F.1) は自明. (S.K.1, 2, 3) ⇒ (F.2) を示す.

p_k $(k=1, 2, \cdots, n)$ がすべて正の有理数のとき，各 p_k は共通の分母によって l_k/m $(m \geqq l_k \geqq 2)$ と表せる. よって (S.K.3) を用いて，

$$S(p) = S(l_1/m, \cdots, l_n/m)$$
$$= S(\underbrace{1/m, \cdots, 1/m}_{l_1\text{ 個}}, \cdots, \underbrace{1/m, \cdots, 1/m}_{l_n\text{ 個}}) - \sum_{k=1}^{n} p_k S(1/l_k, \cdots, 1/l_k).$$

この第一項は変数の分け方の順列 $(l_1, \cdots, l_n)$ に関係しない. また第二項の和 $\sum$ のとり方の順序も任意であるから，$\{p_k\}$ の置換 $\{p_k'\}$ に対して $p_k' = l_k'/m$ とおくと，上式は，

$$S(\underbrace{1/m, \cdots, 1/m}_{l_1'\text{ 個}}, \cdots, \underbrace{1/m, \cdots, 1/m}_{l_n'\text{ 個}}) - \sum p_k' S(1/l_k', \cdots, 1/l_k')$$

に等しい. すなわち，$S(p) = S(p')$. ゆえに (F.2) が成立.

p_k が有理数でないときは，p_k を有理数で近似して，(S.K.1) の連続性を

使えば (F.2) が示せる.

(S.K.1,2,3)⇒(F.3) を示す. (S.K.2,3) と (F.2) より,

$$S(1/2,1/2)=S(1/2,1/2,0,0)=S(1/2,0,1/2,0)$$
$$=S(1/2,1/2)+(1/2)S(1,0)+(1/2)S(1,0).$$

ゆえに, $S(1,0)=0$. したがって,

$$S(p_1,\cdots,p_{n-1},q,r)=S(p_1,0,p_2,0,\cdots,p_{n-1},0,q,r)$$
$$=S(p_1,\cdots,p_n)+\sum_{k=1}^{n-1}p_kS(1,0)+p_nS(q/p_n,r/p_n)$$
$$=S(p_1,\cdots,p_n)+p_nS(q/p_n,r/p_n).$$

よって (F.3) が成立.

[第2ステップ]: (F.1,2,3)⇒$S(p)=-\lambda\sum p_k\log p_k$ を示す.

以下 Tverberg [T. 5] の方法による. (F.2,3) より, 任意の $p,q\geqq 0$, $r>0$, $p+q+r=1$ に対して

$$S(p,q,r)=S(p,q+r)+(q+r)S(q/(q+r),r/(q+r))$$
$$=S(q,p+r)+(p+r)S(p/(p+r),r/(p+r)).$$

ゆえに, $f(p)=S(p,1-p)$ とおくと,

$$f(p)+(1-p)f\left(\frac{q}{1-p}\right)=f(q)+(1-q)f\left(\frac{p}{1-q}\right) \qquad (1.9)$$

となる. (1.9) は任意の $0\leqq p<1$, $0\leqq q<1$ で定義されているので, $q>0$ をとり, $p=0$ とおくと,

$$f(0)=S(0,1)=0$$

を得る. また, (1.9) の両辺を0から $1-p$ まで q に関して積分すると,

$$(1-p)f(p)+(1-p)^2\int_0^1 f(t)dt=\int_0^{1-p}f(t)dt+p^2\int_p^1 t^{-3}f(t)dt \qquad (1.10)$$

となる. 函数 $f(p)$ は連続であるから, (1.10) の左辺第一項を除く項はすべて区間 $0<p<1$ で微分可能である. したがって, $f(p)$ も同じくそうであり, (1.10) の両辺を p で微分すると,

$$(1-p)f'(p)-f(p)-2(1-p)\int_0^1 f(t)dt$$

$$=-f(1-p)+2p\int_p^1 t^{-3}f(t)\,dt-f(p)/p.$$

ここで $f(p)=f(1-p)$ より，

$$(1-p)f'(p)=2(1-p)\int_0^1 f(t)\,dt+2p\int_p^1 t^{-3}f(t)\,dt-f(p)/p \tag{1.11}$$

を得る．(1.11) より，前と同じ議論から $f'(p)$ も $0<p<1$ で微分可能であることがわかる．したがって (1.11) の両辺を p で微分すると，

$$f''(p)=-2\cdot\frac{1}{p(1-p)}\int_0^1 f(t)\,dt \quad (0<p<1) \tag{1.12}$$

となる．(1.12) の両辺を不定積分すると，

$$f(p)=ap+b-2\{p\log p+(1-p)\log(1-p)\}\int_0^1 f(t)\,dt \tag{1.13}$$

が得られる．ここで a, b は積分定数である．ところで $f(p)=f(1-p)$ であるから，$a=0$ となり，さらに (1.13) の両辺は $0\leqq p\leqq 1$ で成立し，$a=0$ と $f(0)=0$ より $b=0$ がいえる．ゆえに，任意の $p_1, p_2\geqq 0$，$p_1+p_2=1$ に対して

$$S(p_1, p_2)=-\lambda(p_1\log p_1+p_2\log p_2)$$

が成立する．ただし，

$$\lambda=2\int_0^1 f(t)\,dt.$$

さて，任意の $p\in\varDelta_n$ に対して，

$$S(p)=-\lambda\sum_{k=1}^{n} p_k\log p_k$$

が成立するとする．このとき，任意の $q\in\varDelta_{n+1}$（ただし，$q_{n+1}>0$ と仮定する．$q_{n+1}=0$ ならば，他の $q_k>0$ のものと q_{n+1} を入れ換えておく）に対しても，

$$S(q)=-\lambda\sum_{k=1}^{n+1} q_k\log q_k$$

となることを示す．

$$\begin{aligned}&S(q_1, q_2, \cdots, q_n, q_{n+1})\\&\quad=S(q_1, \cdots, q_{n-1}, q_n+q_{n+1})\end{aligned}$$

$$+(q_n+q_{n+1})S(q_n/(q_n+q_{n+1}),\ q_{n+1}/(q_n+q_{n+1}))$$
$$=-\lambda\sum_{k=1}^{n-1}q_k\log q_k-\lambda(q_n+q_{n+1})\log(q_n+q_{n+1})$$
$$-\lambda(q_n\log q_n/(q_n+q_{n+1})+q_{n+1}\log q_{n+1}/(q_n+q_{n+1}))$$
$$=-\lambda\sum_{k=1}^{n+1}q_k\log q_k.$$

[第3ステップ]：$S(p)=-\lambda\sum p_k\log p_k$ （$\forall p\in\varDelta_n$） $\Rightarrow$ (S.K.1, 2, 3) を示す.

(S.K.2) は $S(p)$ の形より明らか. また (S.K.3) は〈定理 1.2〉の(3)で証明ずみ. よって (S.K.1) を示せばよい. 任意の $p\in\varDelta_n$ に対して, $q_i=\sum_j a_{ij}p_j$, $a_{ij}=1/n$ ととると, $q=\{1/n\}\in\varDelta_n$. よって〈定理 1.2〉の (2) より,

$$S(p)\leqq S(q)=S\left(\frac{1}{n},\cdots,\frac{1}{n}\right).$$

$S(p)$ の連続性は定義より自明. ∎

最後に, $S(p)$ の主要な性質をまとめておこう. 以下, $p=\{p_k\}$, $q=\{q_j\}$ を $\bigcup_{n\geqq 2}\varDelta_n$ の元, π を添字の置換とする.

〈S.1〉 **正値性**：$S(p)\geqq 0$.

〈S.2〉 **連続性**：$S:\bigcup_{n\geqq 2}\varDelta_n\to\bar{\boldsymbol{R}}^+$ は連続函数.

〈S.3〉 **単調性**：$S\left(\frac{1}{n},\cdots,\frac{1}{n}\right)$ は n に関する増加函数で,

$$S(p_1,\cdots,p_n)\leqq S\left(\frac{1}{n},\cdots,\frac{1}{n}\right),\quad \forall p\in\varDelta_n.$$

〈S.4〉 **拡張可能性**：$S(p_1,\cdots,p_n,0)=S(p_1,\cdots,p_n)$.

〈S.5〉 **対称性**：$S(p_1,\cdots,p_n)=S(p_{\pi(1)},\cdots,p_{\pi(n)})$.

〈S.6〉 **加法性**：$S(p_1q_1,\cdots,p_1q_m,\cdots,p_nq_m)=S(p_1,\cdots,p_n)+S(q_1,\cdots,q_m)$.

〈S.7〉 **劣加法性**：$\sum_{i,j}r_{ij}=1$, $\sum_{j=1}^{m}r_{ij}=p_i$, $\sum_{i=1}^{n}r_{ij}=q_j$ のとき,

$$S(r_{11},\cdots,r_{nm})\leqq S(p_1,\cdots,p_n)+S(q_1,\cdots,q_m).$$

1.5 Shannon エントロピーと Boltzmann エントロピー

この章の最後に Shannon エントロピー $S(p)$ と Boltzmann エントロピー

$k\log W$ の関連を見ておくことにする．

N 個の粒子からなる系を m 個の部分系に分配する仕方，すなわち，全系のとる状態の総数は，縮退度をすべて1とすると，$W_N=N!/N_1!\cdots N_m!$ で与えられた（1.1 節）．ここで $p_k=N_k/N$ とおくと，$p=\{p_k\}_{k=1}^m\in\Delta_m$ であり，〈補題 1.1〉に従うと

$$|(1/N)\log W_N-S_N(p)|=O((1/N)\log N) \tag{1.14}$$

が簡単な計算によりわかる．ここで，$S_N(p)$ は，$S_N(p)=-\sum_{k=1}^m p_k\log p_k$ である．したがって，$N\gg1$ の場合には，1粒子当たりの Boltzmann エントロピーは Shannon エントロピーと近似的に一致することがわかる．

同様に，〈補題 1.1〉を使った計算により次のことがわかる．

$$K_N(p)=\left\{\frac{1}{(2\pi N)^{m-1}\prod_{k=1}^m p_k}\right\}^{1/2}$$

とおくと

$$\left|\frac{K_N(p)}{W_N}e^{NS_N(p)}-1\right|=O\left(\frac{1}{N}\right). \tag{1.15}$$

m 個の部分系がそれぞれ確率 $\{q_k\}_{k=1}^m$ で出現しているとき

$$W_N(p|q)=\frac{N!}{\prod_{k=1}^m(p_kN)!}q_1{}^{p_1N}\cdots q_m{}^{p_mN},$$

$$S_N(p|q)=\sum_{k=1}^m p_k\log(p_k/q_k)$$

とおくと，

$$\left|\frac{K_N(p)}{W_N(p|q)}e^{-NS_N(p|q)}-1\right|=O\left(\frac{1}{N}\right). \tag{1.16}$$

第2章 Kolmogorov-Sinai エントロピー

古典力学における N 個の質点の運動は，位置に関する $3N$ 次元空間 $\boldsymbol{R}^{3N}$ と運動量に関する $3N$ 次元空間 $\boldsymbol{R}^{3N}$ の直積空間，いわゆる相空間 $\Omega=\boldsymbol{R}^{6N}$ において記述される．相空間の点 $\vec{r}=(q,p)\in\Omega$ が Newton (Hamilton) 方程式によって決定される軌道に沿って運動しているとすると，相空間の任意の領域の体積はこの時間発展のもとで不変である (**Liouville の定理**)．さらに，$N\gg 1$ の場合，N 個の粒子すべての初期状態（点）を一意に決めることは不可能であるから，相空間上の初期状態の確率分布とその時間変化を考える必要が生じてくる．

相空間のような連続体の確率分布とその力学的変化を記述するためには，測度論的取り扱いが有効である．以下，この章では一般の確率空間 $(\Omega,\mathscr{F},p)$ の可測有限分割に関するエントロピー (Shannon 流の) を定め，さらに確率空間上の保測変換に関するいわゆる Kolmogorov-Sinai エントロピーに言及し，それらの性質を調べる．

$\mathscr{F}$ を集合 Ω の部分集合族より生成される σ-集合体とし，p を $\mathscr{F}$ 上の確率測度とする．$(\Omega,\mathscr{F},p)$ 上の**保測変換** T とは，$T^{-1}\mathscr{F}\subset\mathscr{F}$ および $p\circ T^{-1}=p$ を満たす Ω から Ω 上への写像のことである．この T が1対1で $T^{-1}\mathscr{F}=\mathscr{F}$ であるとき T は**可逆**であるといい，このとき T^{-1} も保測変換になる．さらに，保測変換である一径数群 $\{T_t;t\in\boldsymbol{R}\}$ $((\omega,t)\in\Omega\times\boldsymbol{R}\to T_t\omega\in\Omega$ が可測) を (保測な) **流れ** (flow) という．以下，我々は4つ組 $(\Omega,\mathscr{F},p,T)$ を基礎として議論を展開する．この4つ組を以下 **"力学系"** とよぶ．

2.1 確率論からの準備

ここで条件付期待値など，以下の議論でしばしば使われる測度論的な確率概

念を簡単に復習しておこう．詳しくは文献［B.7, H.2, I.2, I.7, P.1］などを見よ．

$\mathcal{G}$ を $\mathcal{F}$ の任意の σ-部分集合体として，任意の $f\in L^1(\Omega,\mathcal{F},p)$ $(\equiv\{f:\mathcal{F}$-可測函数; $\|f\|_1=\int_\Omega|f|dp<\infty\}$，これは $L^1(\mathcal{F})$ または $L^1(p)$ と略する場合が多い）に対して

$$\mu(A)=\int_A f\,dp, \qquad \forall A\in\mathcal{G} \tag{2.1}$$

と μ を定めると，これは符号付測度で p に関して絶対連続（i.e., $p(A)=0\Rightarrow\mu(A)=0$）となるから，$p$-a.e. の意味で一意に $\mathcal{G}$-可測函数が存在して，

$$\mu(A)=\int_A g\,dp, \qquad \forall A\in\mathcal{G} \tag{2.2}$$

と表せる（**Radon-Nikodym の定理**による）．この g を f の $\mathcal{G}$ に関する**条件付期待値**といい（以下，単に $\mathcal{G}$-条件付期待値という），$g=E(f|\mathcal{G})$ で表す．この $\mathcal{G}$-条件付期待値は次の性質を有する（p-a.e. の意味で）．

1° $f\geqq 0\Rightarrow E(f|\mathcal{G})\geqq 0$.

2° $E(\cdot|\mathcal{G}):L^1(\mathcal{F})\to L^1(\mathcal{G})$ は線形作用素．

3° $E(E(f|\mathcal{G})|\mathcal{G})=E(f|\mathcal{G})$（射影の性質）．

4° $\mathcal{G}_1\subset\mathcal{G}_2\Rightarrow E(E(f|\mathcal{G}_2)|\mathcal{G}_1)=E(f|\mathcal{G}_1)$.

5° $E(f|\{\emptyset,\Omega\})=\int_\Omega f\,dp$.

6° $f_n\uparrow f\Rightarrow E(f_n|\mathcal{G})\uparrow E(f|\mathcal{G})$.

7° $f\in L^\infty(\mathcal{F})$ $(=\{f:\mathcal{F}$-可測函数; $\|f\|_\infty\equiv\operatorname*{ess\,sup}_{x\in\Omega}|f(x)|=\inf_N(\sup_{x\in\Omega\setminus N}|f(x)|)<\infty\}$，ここで N は p-零集合（i.e., $p(N)=0$, $N\in\mathcal{F}$）である），$g\in L^1(\mathcal{G})$ または $f\in L^1(\mathcal{F})$, $g\in L^\infty(\mathcal{G})$ に対して，

$$E(gf|\mathcal{G})=gE(f|\mathcal{G}).$$

8° $\|E(f|\mathcal{G})\|_1\leqq\|f\|_1$.

逆に，$\mathcal{E}$ を $L^1(\mathcal{F})$ から $L^1(\mathcal{G})$ への写像で，$\mathcal{E}(\mathcal{E}(f)g)=\mathcal{E}(f)g\,(f\in L^1(\mathcal{F}),\ g\in L^\infty(\mathcal{G}))$ および，$\int_A\mathcal{E}(f)dp=\int_A f\,dp$ $(A\in\mathcal{G})$ を満たすとすると，$\mathcal{E}(f)=E(f|\mathcal{G})$ a.e. である．

$\mathcal{G}$-条件付期待値 $E(\cdot|\mathcal{G})$ に関する他の重要な結果を記しておこう．

9° **Jensen** の不等式：$\varphi:\boldsymbol{R}\to\boldsymbol{R}$ が下に有界な凸函数で確率変数 f に対して $\varphi(f)\in L^1(\mathcal{F})$ であれば，$\varphi(E(f|\mathcal{G}))\leqq E(\varphi(f)|\mathcal{G})$.

証明 凸函数は一次函数の sup で表せる．すなわち，

$$\varphi(x)=\sup\{\psi(x);\psi(x)\leqq\varphi(x),\ \psi(x)=\alpha x+\beta \text{ の形}\}.$$

したがって，$\varphi\geqq\psi$ だから

$$E(\varphi(f)|\mathcal{G})\geqq E(\psi(f)|\mathcal{G})=\alpha E(f|\mathcal{G})+\beta.$$

両辺の sup をとると，

$$E(\varphi(f)|\mathcal{G})\geqq\varphi(E(f|\mathcal{G})).\ \blacksquare$$

10° $f\longmapsto E(f|\mathcal{G})$ は $L^p(\mathcal{F})\to L^p(\mathcal{G})$ $(p\geqq1)$ の射影である．

11° $f\longmapsto E(f|\mathcal{G})$ は $L^2(\mathcal{F})\to L^2(\mathcal{G})$ の Hilbert 空間における射影作用素である．つまり，$E(\cdot|\mathcal{G})=P_{L^2(\mathcal{G})}$.

12° $1/p+1/q=1$ のとき，$f\in L^p(\mathcal{F})$，$g\in L^q(\mathcal{F})$ に対して，

$$E(|fg||\mathcal{G})\leqq E(|f|^p|\mathcal{G})^{1/p}E(|g|^q|\mathcal{G})^{1/q}.$$

13° T が保測変換のとき，

$$E(f|\mathcal{G})\circ T=E(f\circ T|T^{-1}\mathcal{G}).$$

ところで，$p(A|\mathcal{G})=E(1_A|\mathcal{G})$ とおくと，これは $\mathcal{G}$ が与えられた場合，事象 $A\in\mathcal{F}$ が生起する確率（条件付確率）であり，以下の性質を有している（p-a.e. の意味で）．

14° $0\leqq p(A|\mathcal{G})\leqq1$.

15° $A\in\mathcal{G}\Rightarrow p(A|\mathcal{G})=1_A$.

16° $A\subset B\Rightarrow p(A|\mathcal{G})\leqq p(B|\mathcal{G})$.

17° $p(A^c|\mathcal{G})=1-p(A|\mathcal{G})$.

18° $p(A|\{\emptyset,\Omega\})=p(A)$.

19° $A\cap B=\emptyset\Rightarrow p(A\cup B|\mathcal{G})=p(A|\mathcal{G})+p(B|\mathcal{G})$.

σ-部分集合体 $\{\mathcal{G}_n\}$ が増大列 $(\mathcal{G}_n\uparrow\mathcal{G})$ であるとき，$f_n\in L^1(\mathcal{G})$ で $E(f_{n+1}|\mathcal{G}_n)=f_n$ を満たす $\mathcal{G}$-可測函数列 $\{f_n\}$ を **martingale** というが，これに関して

20° 収束定理：$\{f_n\}$ が $\{\mathcal{G}_n\}$ に関して martingale であるとき

（1） $\sup\|f_n\|_1<+\infty \Longrightarrow f\in L^1(\mathcal{G})$ が存在して，$f_n\to f$ a.e..

（2） $\{f_n\}$ が一様可積分 (i.e. n に関して一様に $\int_{\{f_n\geq\lambda\}}|f_n|dp\to 0\,(\lambda\to\infty)$)
$\Longrightarrow f\in L^1(\mathcal{G})$ が存在して，$f_n\to f$ a.e. かつ L^1.

21° **増大 martingale**：$\{\mathcal{G}_n\}$ を $\mathcal{F}$ の σ-部分集合体の増大列 $\mathcal{G}_n\uparrow\mathcal{G}$ (i.e., $\mathcal{G}=\bigvee_n\mathcal{G}_n$) とする．このとき $f\in L^1(\mathcal{F})$ に対して，$E(f|\mathcal{G}_n)\to E(f|\mathcal{G})$ a.e. かつ L^1.

22° **減少 martingale**：$\{\mathcal{G}_n\}$ を $\mathcal{F}$ の σ-部分集合体の減少列 $\mathcal{G}_n\downarrow\mathcal{G}$ (i.e., $\mathcal{G}=\bigcap_n\mathcal{G}_n$) とする．このとき $f\in L^1(\mathcal{F})$ に対して，$E(f|\mathcal{G}_n)\to E(f|\mathcal{G})$ a.e. かつ L^1.

ここで 21° を証明しておこう．詳しくは [B.7, P.1] などを参照していただきたい．

$\mathcal{G}=\mathcal{F}$ としても一般性を失わない．よって $E(f|\mathcal{G})=f$．任意の $f\in L^1(\mathcal{F})$ と $\varepsilon>0$ に対して，$\bigcup_n L^1(\Omega,\mathcal{G}_n,p)$ は $L^1(\Omega,\mathcal{F},p)$ で稠密だから，ある番号 k と $g\in L^1(\Omega,\mathcal{G}_k,p)$ が存在して，

$$\int|f-g|dp<\varepsilon.$$

$n\geq k$ に対して，$\mathcal{G}_n\supset\mathcal{G}_k$ であるから，$E(g|\mathcal{G}_n)=g$ であり，したがって，

$$\begin{aligned}&\int|E(f|\mathcal{G}_n)-f|dp\\&\leq\int|E(f|\mathcal{G}_n)-E(g|\mathcal{G}_n)|dp+\int|E(g|\mathcal{G}_n)-g|dp+\int|g-f|dp\\&\leq 2\int|g-f|dp<2\varepsilon.\end{aligned}$$

$\varepsilon>0$ は任意だから，L^1-ノルムで，$E(f|\mathcal{G}_n)\to f$．次に a.e. 収束が次のようにして示される．

$$\begin{aligned}&p\{\omega\in\Omega;\ \limsup|E(f|\mathcal{G}_n)-f|>\varepsilon^{1/2}\}\\&\leq p\{\omega\in\Omega;\ \limsup[|E((f-g)|\mathcal{G}_n)-(f-g)|+|E(g|\mathcal{G}_n)-g|]>\varepsilon^{1/2}\}\\&\leq p\{\omega\in\Omega;\ \limsup[|E((f-g)|\mathcal{G}_n)|+|f-g|]>\varepsilon^{1/2}\}\end{aligned}$$

$$\leqq p\{\omega\in\Omega;\ \limsup|E((f-g)|\mathcal{G}_n)|>(1/2)\varepsilon^{1/2}\}$$
$$+p\{\omega\in\Omega;|f-g|>(1/2)\varepsilon^{1/2}\}$$
$$\leqq(2/\varepsilon^{1/2})\int|f-g|dp+(2/\varepsilon^{1/2})\int|f-g|dp\leqq 4\varepsilon^{1/2}.$$

$\varepsilon>0$ は任意だから，$E(f|\mathcal{G}_n)\xrightarrow{\text{a.e.}}f$. ▮

2.2 有限分割のエントロピー

可測空間 $(\Omega,\mathcal{F})$ と $\mathcal{F}$ の σ-部分集合体 $\mathcal{G}$ において，$\tilde{A}=\{A_1,A_2,\cdots,A_n\}$ が Ω の**可測有限分割**（以下，略して $\mathcal{G}$-分割という）であるとは，$A_k\in\mathcal{G}$, $A_l\cap A_k=\phi$, $\bigcup_{k=1}^{n}A_k=\Omega$ となる場合をいう．$\mathcal{G}$-分割の全体を $\mathcal{P}(\mathcal{G})$ で表す．任意の $\tilde{A},\tilde{B}\in\mathcal{P}(\mathcal{G})$ に対して，

$$\tilde{A}\vee\tilde{B}=\{A_l\cap B_j;A_l\in\tilde{A},B_j\in\tilde{B}\},$$
$$T^{-1}\tilde{A}=\{T^{-1}A;A\in\tilde{A}\}$$

と定めると，これらも $\mathcal{G}$-分割である．

ここでは $\tilde{A}$ によって生成される σ-集合体（有限集合体）も同じ記号 $\tilde{A}$ によって表すことにする．したがって，$\bigvee_{k=1}^{n}\tilde{A}_k$ は $\bigcup_{k=1}^{n}\tilde{A}_k$ によって生成される σ-集合体である．ところで，σ-集合体 $\mathcal{F}$ の元 A が $\mathcal{F}$ のアトム (atom) であるとは $B\subset A$, $p(B)<p(A)$ なる $B\in\mathcal{F}$ に対してつねに $p(B)=0$ となる場合であるが，$\tilde{A}$ の元 A_k は $\tilde{A}$ によって生成される σ-集合体のアトムになっている．

任意の $\tilde{A}\in\mathcal{P}(\mathcal{F})$ に対して，分割のエントロピーは

$$S(\tilde{A})=-\sum_{k=1}^{n}p(A_k)\log p(A_k). \tag{2.3}$$

以下，このエントロピーは $\tilde{A}$ のアトムについて和をとるという意味で，

$$S(\tilde{A})=-\sum_{A\in\tilde{A}}p(A)\log p(A)$$

とかく．また，函数

$$\eta(t)=\begin{cases}-t\log t & (0<t\leqq 1)\\ 0 & (t=0)\end{cases}$$

を使い，

$$S(\tilde{A})=\sum_{k=1}^{n}\eta(p(A_k))\quad (=\sum_{A\in\tilde{A}}\eta(p(A)))$$

と表すことも多い．さらに，$1_A(\omega)$ を特性函数とするとき

$$I(\tilde{A})(\omega)=-\sum_{A\in\tilde{A}}\log p(A)1_A(\omega),\qquad \forall\omega\in\Omega$$

を $\tilde{A}$ の**エントロピー函数**という．このとき，

$$S(\tilde{A})=E(I(\tilde{A}))=\int_\Omega I(\tilde{A})dp.$$

$\mathcal{G}$ を $\mathcal{F}$ の部分集合体とするとき，**条件付エントロピー函数**，**条件付エントロピー**を次のように定義する．

$$I(\tilde{A}|\mathcal{G})(\omega)=-\sum_{A\in\tilde{A}}\log p(A|\mathcal{G})1_A(\omega),$$

$$S(\tilde{A}|\mathcal{G})=E(I(\tilde{A}|\mathcal{G}))=\int_\Omega I(\tilde{A}|\mathcal{G})dp=-\sum_{A\in\tilde{A}}\int_A\log p(A|\mathcal{G})dp.$$

したがって，$\tilde{A},\tilde{B}\in\mathcal{P}(\mathcal{F})$ に対して，

$$\begin{aligned}S(\tilde{A}|\tilde{B})=\int_\Omega I(\tilde{A}|\tilde{B})dp&=-\sum_{A\in\tilde{A}}\int_A\sum_{B\in\tilde{B}}\log(p(A|B)1_B)dp\\&=-\sum_{A,B}p(A\cap B)\log p(A|B)\\&=\sum_{B\in\tilde{B}}p(B)\sum_{A\in\tilde{A}}\{-p(A|B)\log p(A|B)\}\end{aligned}\tag{2.4}$$

となる．

以下，$\Omega,\emptyset$ からなる有限集合体（$\mathcal{F}$-分割）を $\tilde{\mathcal{O}}=\{\Omega,\emptyset\}$ で表すことにする．

〈定理 2.1〉 任意の $\tilde{A},\tilde{B}\in\mathcal{P}(\mathcal{F})$ と $\mathcal{F}$ の任意の部分集合体 $\mathcal{G},\mathcal{G}_k$ ($k=1,2$) および保測変換 T に対して

（1） $S(\tilde{A}|\tilde{\mathcal{O}})=S(\tilde{A})$.

（2） $S(\tilde{A}\vee\tilde{B}|\mathcal{G})=S(\tilde{A}|\mathcal{G})+S(\tilde{B}|\tilde{A}\vee\mathcal{G})$.

（2′） $S(\tilde{A}\vee\tilde{B})=S(\tilde{A})+S(\tilde{B}|\tilde{A})$.

（3）　$\tilde{A}\subset\tilde{B}\Rightarrow S(\tilde{A}|\mathcal{G})\leqq S(\tilde{B}|\mathcal{G})$.

（3′）　$\tilde{A}\subset\tilde{B}\Rightarrow S(\tilde{A})\leqq S(\tilde{B})$.

（4）　$\mathcal{G}_1\supset\mathcal{G}_2\Rightarrow S(\tilde{A}|\mathcal{G}_1)\leqq S(\tilde{A}|\mathcal{G}_2)$.

（4′）　$S(\tilde{A}|\mathcal{G})\leqq S(\tilde{A})$.

（5）　$S(\tilde{A}\vee\tilde{B}|\mathcal{G})\leqq S(\tilde{A}|\mathcal{G})+S(\tilde{B}|\mathcal{G})$.

（5′）　$S(\tilde{A}\vee\tilde{B})\leqq S(\tilde{A})+S(\tilde{B})$.

（6）　$S(T^{-1}\tilde{A}|T^{-1}\mathcal{G})=S(\tilde{A}|\mathcal{G})$.

（6′）　$S(T^{-1}\tilde{A})=S(\tilde{A})$.

証明　証明のほとんどは，第1章の $S(X)$ の場合と同様である．

（1）：$S(\tilde{A}/\tilde{\mathcal{O}})=\sum_{A\in\tilde{A}}\{-p(A\cap\Omega)\log p(A|\Omega)-p(A\cap\emptyset)\log p(A|\emptyset)\}$

$$=-\sum_{A\in\tilde{A}}p(A)\log p(A)=S(\tilde{A}).$$

この（1）の等式より，命題 (n') は命題 (n) $(n=1,2,\cdots,6)$ が示されれば，$\tilde{B}=\tilde{\mathcal{O}}$ または $\mathcal{G}=\tilde{\mathcal{O}}$ $(\mathcal{G}_k=\tilde{\mathcal{O}})$ とおくことにより得られる．

（2）：$\tilde{A}$ は有限集合体であるから，

$$p(B|\tilde{A}\vee\mathcal{G})=\sum_{A\in\tilde{A}}1_A\cdot\frac{p(B\cap A|\mathcal{G})}{p(A|\mathcal{G})}\quad\text{a.e.}$$

である．よって，

$$\log p(B|\tilde{A}\vee\mathcal{G})=\sum_{A\in\tilde{A}}1_A\cdot\log\frac{p(A\cap B|\mathcal{G})}{p(A|\mathcal{G})}\quad\text{a.e.}.$$

ゆえに，

$$\begin{aligned}&\sum_{A,B}1_{A\cap B}\cdot\log p(A\cap B|\mathcal{G})\\&=\sum_{A,B}1_A\cdot 1_B\cdot\log p(A|\mathcal{G})+\sum_{A,B}1_A\cdot 1_B\cdot\log(p(A\cap B|\mathcal{G})/p(A|\mathcal{G}))\\&=\sum_A 1_A\cdot\log p(A|\mathcal{G})+\sum_B 1_B\log p(B|\tilde{A}\vee\mathcal{G}).\end{aligned}$$

この両辺の期待値をとると，$S(\tilde{A}|\mathcal{G})$ の定義より（2）が得られる．

（3）：$\tilde{A}\subset\tilde{B}$であるから，$\tilde{A}\vee\tilde{B}=\tilde{B}$ となる．よって（2）から明らか．

（4）：$\eta(t)$ は凹函数であるから Jensen の不等式より

$$E(\eta(p(A|\mathcal{G}_1))|\mathcal{G}_2)\leqq\eta(E(p(A|\mathcal{G}_1)|\mathcal{G}_2)),$$

条件付期待値の性質より

$$E(p(A|\mathcal{G}_1)|\mathcal{G}_2)=E(E(1_A|\mathcal{G}_1)|\mathcal{G}_2)=E(1_A|\mathcal{G}_2)=p(A|\mathcal{G}_2).$$

よって

$$E(\eta(p(A|\mathcal{G}_1))|\mathcal{G}_2)\leqq\eta(p(A|\mathcal{G}_2)).$$

両辺の期待値をとり，$\tilde{A}$ のアトムに関して和をとると

$$S(\tilde{A}|\mathcal{G}_1)=\sum_{A} E(\eta(p(A|\mathcal{G}_1)))\leqq\sum_{A} E(\eta(p(A|\mathcal{G}_2)))=S(\tilde{A}|\mathcal{G}_2).$$

（5）：$\tilde{A}\vee\mathcal{G}\supset\mathcal{G}$ であるから，（4）より，$S(\tilde{B}|\tilde{A}\vee\mathcal{G})\leqq S(\tilde{B}|\mathcal{G})$. したがって（2）より，

$$S(\tilde{A}\vee\tilde{B}|\mathcal{G})=S(\tilde{A}|\mathcal{G})+S(\tilde{B}|\tilde{A}\vee\mathcal{G})\leqq S(\tilde{A}|\mathcal{G})+S(\tilde{B}|\mathcal{G}).$$

（6）：$p(T^{-1}A|T^{-1}\mathcal{G})(\omega)=p(A|\mathcal{G})(T\omega)$ a.e.，$\omega\in\Omega$ であるから，

$$E(\eta(p(T^{-1}A|T^{-1}\mathcal{G})))=E(\eta(p(A|\mathcal{G})))$$

を得る．ここで $A\in\tilde{A}$ について加え合わせればよい．▌

〈系 2.2〉〈定理 2.1〉におけるすべての関係式は，$I(\tilde{A})$ および $I(\tilde{A}|\mathcal{G})$ においても成立する．ただし（6），(6′) は次の等式に置き換わる．

（6）　$I(T^{-1}\tilde{A})=I(\tilde{A})\circ T.$

(6′)　$I(T^{-1}\tilde{A}|T^{-1}\mathcal{G})=I(\tilde{A}|\mathcal{G})\circ T.$

この証明は読者の演習問題としよう．

註　上記の定理 (2′) において $\tilde{B}\subset\tilde{A}$ ととると，$S(\tilde{B}|\tilde{A})=0$ を得る．よって $\tilde{A}\subset\mathcal{G}$ のとき $S(\tilde{A}|\mathcal{G})=0$ となる．

2.3 Kolmogorov-Sinai エントロピー

有限個の事象をもつ系（実験）$\tilde{A}$ を n 回繰り返して $\tilde{A}$, $T^{-1}\tilde{A}$, $T^{-2}\tilde{A},\cdots$, $T^{-(n-1)}\tilde{A}$ を得たとき，合成系のエントロピー $S\left(\bigvee_{k=0}^{n-1}T^{-k}\tilde{A}\right)$ から $\tilde{A}, T$ に関する実験1回当たりのエントロピー $S(\tilde{A},T)$ は

$$S(\tilde{A},T)\equiv\lim_{n\to\infty}\frac{1}{n}S\left(\bigvee_{k=0}^{n-1}T^{-k}\tilde{A}\right)\tag{2.5}$$

で与えられる（この極限の存在は〈補題 2.4〉の（1）で示す）．そこで，

$$S(T)=\sup\{S(\tilde{A},T);\ \tilde{A}\in\mathcal{P}(\mathcal{F})\} \tag{2.6}$$

とおき，この $S(T)$ を**保測変換 T のエントロピー**，または，**Kolmogorov-Sinai エントロピー**という．

以下の補題，定理等において，とくに断らない限り，T は保測変換，$\tilde{A},\tilde{B}$ 等は有限分割とする．

〈補題 2.3〉 次の等式が成立する．

（1）　$S\left(\bigvee_{k=0}^{n-1}T^{-k}\tilde{A}\right)=S(\tilde{A})+\sum_{i=1}^{n-1}S\left(\tilde{A}|\bigvee_{k=1}^{i}T^{-k}\tilde{A}\right).$

（2）　$I\left(\bigvee_{k=0}^{n-1}T^{-k}\tilde{A}\right)=I(\tilde{A})\circ T^{n-1}+\sum_{i=1}^{n-1}I\left(\tilde{A}|\bigvee_{k=1}^{i}T^{-k}\tilde{A}\right).$

証明　〈定理 2.1〉の (2′) および (6′) より（1）は自明．（2）は $I(\tilde{A})$ に対する (2′) および (6′) より求まる．▮

〈補題 2.4〉 次の等式が成立する．

（1）　$S(\tilde{A},T)=\lim_{n\to\infty}S\left(\tilde{A}|\bigvee_{k=1}^{n}T^{-k}\tilde{A}\right).$

（2）　$S(\tilde{A},T)=\lim_{n\to\infty}S\left(T^{-n}\tilde{A}|\bigvee_{k=0}^{n-1}T^{-k}\tilde{A}\right).$

（3）　T が可逆なとき，$S(\tilde{A},T)=\lim_{n\to\infty}S\left(\tilde{A}|\bigvee_{k=1}^{n}T^{k}\tilde{A}\right).$

証明　（1）:〈定理 2.1〉の（4）より

$$S\left(\tilde{A}|\bigvee_{k=1}^{i}T^{-k}\tilde{A}\right)\leq S\left(\tilde{A}|\bigvee_{k=1}^{i-1}T^{-k}\tilde{A}\right).$$

よって〈補題 2.3〉の（1）より，

$$S(\tilde{A},T)=\lim_{n\to\infty}(1/n)S\left(\bigvee_{k=0}^{n-1}T^{-k}\tilde{A}\right)=\lim_{n\to\infty}(1/n)\sum_{i=1}^{n-1}S\left(\tilde{A}|\bigvee_{k=1}^{i}T^{-k}\tilde{A}\right)$$
$$=\lim_{n\to\infty}S\left(\tilde{A}|\bigvee_{k=1}^{n}T^{-k}\tilde{A}\right).$$

（2）:〈定理 2.1〉の（6）より

$$S\left(T^{-i}\tilde{A}|\bigvee_{k=0}^{i-1}T^{-k}\tilde{A}\right)=S\left(T^{-(i+1)}\tilde{A}|\bigvee_{k=1}^{i}T^{-k}\tilde{A}\right)\geqq S\left(T^{-(i+1)}\tilde{A}|\bigvee_{k=0}^{i}T^{-k}\tilde{A}\right).$$

および〈定理 2.1〉の (2′) より

$$S\left(\bigvee_{k=0}^{n}T^{-k}\tilde{A}\right)=S(\tilde{A})+\sum_{i=1}^{n}S\left(T^{-i}\tilde{A}|\bigvee_{k=0}^{i-1}T^{-k}\tilde{A}\right).$$

したがって,

$$S(\tilde{A},T)=\lim_{n\to\infty}(1/n)S\left(\bigvee_{k=0}^{n}T^{-k}\tilde{A}\right)=\lim_{n\to\infty}(1/n)\sum_{i=1}^{n}S\left(T^{-i}\tilde{A}|\bigvee_{k=0}^{i-1}T^{-k}\tilde{A}\right)$$

$$=\lim_{n\to\infty}S\left(T^{-n}\tilde{A}|\bigvee_{k=0}^{n-1}T^{-k}\tilde{A}\right).$$

（3）:〈定理 2.1〉の（6）より

$$S\left(\bigvee_{k=1}^{n}T^{k}\tilde{A}\right)=S\left(T^{-(n-1)}\bigvee_{k=1}^{n}T^{k}\tilde{A}\right)=S\left(\bigvee_{k=1}^{n}T^{-k}\tilde{A}\right).$$

よって（3）が成立. ▮

〈定理 2.5〉 次の等式が成立する.

（1） $\tilde{A}\subset\tilde{B}\Longrightarrow S(\tilde{A},T)\leqq S(\tilde{B},T)$.

（2） $S\left(\bigvee_{k=m}^{n}T^{-k}\tilde{A},T\right)=S(\tilde{A},T)\quad(m\geqq 0)$.

等式（2）は T が可逆ならば, 任意の整数 m において成立.

（3） $S\left(\bigvee_{k=0}^{m-1}T^{-k}A,T^{m}\right)=mS(\tilde{A},T)\quad(m\geqq 1)$.

（4） $S(\tilde{A},T)\leqq S(\tilde{B},T)+S(\tilde{A}|\tilde{B})$.

証明 （1）: 定義式 (2.5) と〈定理 2.1〉の (3′) より自明.

（2）: $t\geqq m\geqq 0$ に対して

$$\bigvee_{k=0}^{n-1}T^{-k}\left\{\bigvee_{j=m}^{t}T^{-j}\tilde{A}\right\}=T^{-m}\left\{\bigvee_{k=0}^{n+t-m-1}T^{-k}\tilde{A}\right\}.$$

〈定理 2.1〉の (6′) より

$$\frac{1}{n}S\left(\bigvee_{k=0}^{n-1}T^{-k}\left\{\bigvee_{j=m}^{t}T^{-j}\tilde{A}\right\}\right)=\frac{n+t-m}{n}\,\frac{1}{n+t-m}S\left(\bigvee_{k=0}^{n+t-m-1}T^{-k}\tilde{A}\right).$$

$n\to\infty$ をとると,

$$S\left(\bigvee_{j=m}^{l} T^{-j}\tilde{A}, T\right)=S(\tilde{A}, T).$$

（3）：$m\geqq 1$ であるから，

$$\frac{1}{n}S\left(\bigvee_{k=0}^{n-1}(T^m)^{-k}\left\{\bigvee_{j=0}^{m-1}T^{-j}\tilde{A}\right\}\right)=m\cdot\frac{1}{nm}S\left(\bigvee_{j=0}^{nm-1}T^{-j}\tilde{A}\right).$$

$n\to\infty$ をとると，

$$S\left(\bigvee_{j=0}^{m-1}T^{-j}\tilde{A}, T^m\right)=mS(\tilde{A}, T).$$

（4）：〈定理 2.1〉の (2′), (3′) より，

$$\begin{aligned}S\left(\bigvee_{k=0}^{n-1}T^{-k}\tilde{A}\right)&\leqq S\left(\left\{\bigvee_{k=0}^{n-1}T^{-k}\tilde{A}\right\}\vee\left\{\bigvee_{j=0}^{n-1}T^{-j}\tilde{B}\right\}\right)\\&=S\left(\bigvee_{k=0}^{n-1}T^{-k}\tilde{B}\right)+S\left(\bigvee_{k=0}^{n-1}T^{-k}\tilde{A}\,\middle|\,\bigvee_{j=0}^{n-1}T^{-j}\tilde{B}\right)\\&\leqq S\left(\bigvee_{k=0}^{n-1}T^{-k}\tilde{B}\right)+\sum_{k=0}^{n-1}S\left(T^{-k}\tilde{A}\,\middle|\,\bigvee_{j=0}^{n-1}T^{-j}\tilde{B}\right)\\&\leqq S\left(\bigvee_{k=0}^{n-1}T^{-k}\tilde{B}\right)+\sum_{k=0}^{n-1}S(T^{-k}\tilde{A}|T^{-k}\tilde{B})\\&=S\left(\bigvee_{k=0}^{n-1}T^{-k}\tilde{B}\right)+nS(\tilde{A}|\tilde{B}).\end{aligned}$$

ここで，両辺を n で割った後 $n\to\infty$ ととると，

$$S(\tilde{A}, T)\leqq S(\tilde{B}, T)+S(\tilde{A}|\tilde{B}).\ \blacksquare$$

この定理の（2）と（3）より次の系が容易に示される．

〈系 2.6〉 $m\geqq 1$ のとき，$S(T^m)=mS(T)$．さらに，T が可逆のときはすべての整数 m に対して，$S(T^m)=mS(T)$ が成立する．

註 ここで考えている確率測度空間 $(\Omega, \mathcal{F}, p)$ が Lebesgue 空間（すなわち，区間 [0, 1] に高々可算個の点が付いた空間として表現される空間 [W. 1]）のとき，流れ $\{T_t; t\in\boldsymbol{R}\}$ に対して，

$$S(T_t)=|t|S(T_1),\qquad \forall t\in\boldsymbol{R}$$

が成立することが，Abramov [A. 1] によって示された．この証明は Totoki [T. 4] などを参照されたい．

ところで，保測変換 T_1, T_2 をもつ2つの力学系 $(X_1, \mathcal{F}_1, p_1, T_1)$ と $(X_2,$

$\mathcal{F}_2, p_2, T_2)$ が同型であるとは，X_1 から X_2 の上への写像 φ が存在して，(i) φ は 1 対 1；(ii) $A_2=\varphi(A_1)$ のとき $A_1\in\mathcal{F}_1 \Longleftrightarrow A_2\in\mathcal{F}_2$，かつ $p_1(A_1)=p_2(A_2)$；(iii) $\varphi\circ T_1=T_2\circ\varphi$ を満たすときをいい，$T_1\cong T_2$ で表す．Kolmogorov-Sinai エントロピー $S(T)$ はこの同型変換に対して不変量である．

実際，T_1 と T_2 が同型であるとし，φ を同型写像とすると，φ の性質 (i) ～(iii) より

$$S(\tilde{A}_2, T_2)=S(\varphi(\tilde{A}_1), T_1)$$

となる．ゆえに，$S(T_2)\leqq S(T_1)$．同様に $S(T_1)\leqq S(T_2)$ もわかるので，$S(T_1)=S(T_2)$ となる．

今，$S(T)=0$ であるとすると，$S(T)$ の定義より $S(\tilde{A}, T)=0$ $(\forall\tilde{A}\in\mathcal{P}(\mathcal{F}))$ となるが，このことより $S(T)=0$ であれば $T^{-1}\mathcal{F}=\mathcal{F}$ となる．実際，$T^{-1}\mathcal{F}\subsetneqq\mathcal{F}$ とすると，$B\notin T^{-1}\mathcal{F}$ である $B\in\mathcal{F}$ が存在する．ここで $\tilde{A}=\{B, B^c\}$ ととると，$S(\tilde{A}|T^{-1}\mathcal{F})\neq 0$ より，$S(\tilde{A}, T)=\lim_{n\to\infty}S\left(\tilde{A}|\bigvee_{k=0}^{n-1}T^{-k}\tilde{A}\right)\neq 0$ となるから，$S(T)\neq 0$.

〈定理 2.7〉 2つの保測変換 T_1, T_2 が同型ならば，

$$S(T_1)=S(T_2).$$

さて，$S(T)$ のある意味での計算方法を与える Kolmogorov-Sinai の定理を証明しよう．まず，有用な 2 つの補題を証明する．

〈補題 2.8〉 $\mathcal{F}$ の σ-部分集合体の増大列 $\mathcal{G}_n\uparrow\mathcal{G}$ に対して，

$$\int_\Omega \sup_n I(\tilde{A}|\mathcal{G}_n)\,dp\leqq S(\tilde{A})+1.$$

証明 $f=\sup_n I(\tilde{A}|\mathcal{G}_n)$，$F(t)=p\{\omega\in\Omega; f(\omega)>t\}$ とおくと，

$$\int f\,dp=-\int_0^\infty t\,dF(t)=[-tF(t)]_0^\infty+\int_0^\infty F(t)dt\leqq\int_0^\infty F(t)dt,$$

$$F(t)=p\{\omega\in\Omega;\ \sup_n(-\sum_{A\in\tilde{A}}1_A\log p(A|\mathcal{G}_n))>t\}$$

$$= \sum_{A \in \tilde{A}} p(A \cap \{\omega \in \Omega;\ \inf_n p(A|\mathcal{G}_n) < e^{-t}\}).$$

ところで

$$p(A \cap \{\omega \in \Omega;\ \inf_n p(A|\mathcal{G}_n) < e^{-t}\})$$

$$= \sum_{n=1}^{\infty} p(A \cap A_n) = \sum_{n=1}^{\infty} \int_{A_n} p(A|\mathcal{G}_n) dp.$$

ここで，$A_n = \bigcap_{k<n} \{\omega \in \Omega;\ p(A|\mathcal{G}_n) < e^{-t},\ p(A|\mathcal{G}_k) > e^{-t}\}$.

$$\int_{A_n} p(A|\mathcal{G}_n) dp < e^{-t} p(A_n)$$

であるから，

$$\sum_{n=1}^{\infty} p(A \cap A_n) \leqq e^{-t}.$$

よって

$$F(t) \leqq \sum_{A \in \tilde{A}} \min(p(A), e^{-t}).$$

ゆえに

$$\int_0^\infty F(t) dt \leqq \sum_{A \in \tilde{A}} \int_0^\infty \min(p(A), e^{-t}) dt = \sum_{A \in \tilde{A}} (-p(A) \log p(A) + p(A))$$

$$= S(\tilde{A}) + 1. \quad \blacksquare$$

〈**補題 2.9**〉 $\mathcal{F}$ の σ-部分集合体の増大列 $\mathcal{G}_n \uparrow \mathcal{G}$ に対して，

（1） $I(\tilde{A}|\mathcal{G}_n) \to I(\tilde{A}|\mathcal{G})$ a.e. かつ L^1-収束；

（2） $S(\tilde{A}|\mathcal{G}_n) \downarrow S(\tilde{A}|\mathcal{G})$.

証明 各$A \in \tilde{A}$ に対して martingale の収束定理（第2章 2.1 節の (21)）より，$p(A|\mathcal{G}_n) \to p(A|\mathcal{G})$ a.e.. 前補題より $I(\tilde{A}|\mathcal{G}_n)$ は可積分函数で押さえられるから，$L^1(\Omega)$ においても $I(\tilde{A}|\mathcal{G}_n) \to I(\tilde{A}|\mathcal{G})$ である．また

$$S(\tilde{A}|\mathcal{G}_n) = \int_\Omega I(\tilde{A}|\mathcal{G}_n) dp$$

であり，$S(\tilde{A}|\mathcal{G}_n) \geqq S(\tilde{A}|\mathcal{G}_{n+1})$ であるから，

$$S(\tilde{A}|\mathcal{G}_n) \downarrow S(\tilde{A}|\mathcal{G}). \quad \blacksquare$$

これらの補題を使うと，

〈定理 2.10〉 Kolmogorov-Sinai の定理：T が可逆で，$\tilde{A}\in\mathcal{P}(\mathcal{F})$ に対して，$\bigvee_{n=-\infty}^{\infty} T^n\tilde{A}=\mathcal{F}$ であれば，$S(T)=S(\tilde{A}, T)$.

証明 $\tilde{A}^{(n)}=\bigvee_{k=-n}^{n} T^k\tilde{A}$ とおくと，〈定理 2.5〉の（2）より，

$$S(\tilde{A}^{(n)}, T)=S(\tilde{A}, T).$$

また，〈定理 2.5〉の（4）より任意の $\tilde{B}\in\mathcal{P}(\mathcal{F})$ に対して，

$$\begin{aligned} S(\tilde{B}, T) &\leqq S(\tilde{A}^{(n)}, T)+S(\tilde{B}|\tilde{A}^{(n)}) \\ &=S(\tilde{A}, T)+S(\tilde{B}|\tilde{A}^{(n)}). \end{aligned}$$

したがって，$\lim_{n\to\infty} S(\tilde{B}|\tilde{A}^{(n)})=0$ が成立すれば，

$$S(T)=\sup\{S(\tilde{B}, T);\ \tilde{B}\in\mathcal{P}(\mathcal{F})\}\leqq S(\tilde{A}, T)\leqq S(T)$$

より証明が終わる．ところが，$\tilde{A}^{(n)}\uparrow\mathcal{F}$ であるから〈補題 2.9〉より，

$$\lim_{n\to\infty} S(\tilde{B}|\tilde{A}^{(n)})=S(\tilde{B}|\mathcal{F})=0. \ \blacksquare$$

T は必ずしも可逆ではないが，そのときは，

〈系 2.11〉 $\bigvee_{n=0}^{\infty} T^{-n}\tilde{A}=\mathcal{F}$ ならば，$S(T)=S(\tilde{A}, T)$ が成立する．

この証明は〈定理 2.10〉の証明において $\tilde{A}^{(n)}=\bigvee_{k=0}^{\infty} T^{-k}\tilde{A}$ とおけばよい．さらに，

〈系 2.12〉 T が可逆で，$\bigvee_{n=0}^{\infty} T^{-n}\tilde{A}=\mathcal{F}$ ならば，

$$S(T)=S(\tilde{A}, T)=0.$$

証明 T が可逆であるから，

$$\tilde{A}\subset\mathcal{F}=T^{-1}\mathcal{F}=\bigvee_{n=1}^{\infty} T^{-n}\tilde{A}.$$

ゆえに，〈補題 2.9〉より

$$\lim_{n\to\infty} S\left(\tilde{A}\middle|\bigvee_{k=1}^{n} T^{-k}\tilde{A}\right)=S(\tilde{A}|\mathcal{F})=0.$$

したがって，〈補題 2.4〉の（1）より，$S(\tilde{A}, T)=0$. よって，$S(T)=0$. $\blacksquare$

2つの力学系が同型であるとき，すでに議論したようにエントロピー $S(T)$ は不変量であった．したがって，エントロピー $S(T)$ が不変量でない2つの力学系は同型ではないことがわかる．

第5章で詳細に議論するが，ある有限集合（alphabet の集合など）A からつくられる確率空間$(A^Z, \mathcal{F}, p)$とその上のシフト$T: A^Z \to A^Z$ $(T\Pi x_k = \Pi x_{k+1}$, $x_k \in A)$ を考える．$P^{(n)} = (P_{ij}{}^{(n)})$ を時刻nでの既約な $k \times k$ 確率行列とする（i.e., $\sum_{j=1}^{k} P_{ij}{}^{(n)} = 1$, かつ任意の i, j に対して $m \in N$ が存在して $P_{ij}{}^{(m)} > 0$)．各筒集合 $[i_0 i_1 \cdots i_l] \equiv \{\prod_k x_k \in A^Z;\ x_k = i_k,\ 0 \leq k \leq l\}$ に対して，p が

$$p([i_0 i_1 \cdots i_l]) = P_{i_0} P_{i_0 i_1} \cdots P_{i_{l-1} i_l} \quad (P_{ij} \equiv P_{ij}{}^{(1)})$$

で与えられるとき，とくに $P_{ij} = p_i$ であり p が上式で与えられているとき，T を **Bernoulli シフト**という．ここで p_i は $i \in A$ が起こる確率である．さらに，$\{p_i\}$ と Bernoulli シフトTをもつ変換を，(p_i)**-Bernoulli 変換**という．$\tilde{A} = ([x_0 = i])_{i=1,\cdots,|A|}$ とおくと $\bigvee_{n=-\infty}^{\infty} T^n \tilde{A} = \mathcal{F}$ であるから，〈定理 2.10〉より，

$$S(T) = S(\tilde{A}, T) = \lim_{n\to\infty} S\left(\tilde{A} \middle| \bigvee_{k=1}^{n} T^k \tilde{A}\right).$$

$\bigvee_{k=1}^{n} T^k \tilde{A}$ のアトムは $[i_{-1} \cdots i_{-n}]$ であるから

$$S\left(\tilde{A} \middle| \bigvee_{k=1}^{n} T^k \tilde{A}\right) = \sum_{i_{-1},\cdots,i_{-n}} p([i_{-1} \cdots i_{-n}]),$$

$$\sum_{i_0} \eta(p(x_0 = i_0;\ x_{-1} = i_{-1}, \cdots, x_{-n} = i_{-n})).$$

(p_i)-Bernoulli 変換では，$p(i_0 | i_{-1}, \cdots, i_{-n}) = p_{i_0}$ だから，$S(T) = -\sum_i p_i \log p_i$ となる．したがって (1/2, 1/2)-Bernoulli 変換と (1/3, 1/3, 1/3)-Bernoulli 変換は同型でないことがわかる．

では，〈定理 2.7〉の逆，すなわち，“エントロピーの等しい2つの保測変換は同型である”は正しいであろうか？ この問は，Meshalkin, Sinai の仕事を経て，Ornstein によって次の形で答えられている．

〈定理 2.13〉 エントロピーの等しい2つの Bernoulli 変換は同型である．

この定理の証明は，Totoki [T.4] あるいは Ornstein の原論文 [O.5] を参照されたい．また，Ornstein と Shields [O.6] によって示されているように，$S(T_1)=S(T_2)$ であるが $T_1 \not\cong T_2$ となる保測変換は非可算個存在する．よって，エントロピーは Bernoulli 変換に関して完全不変量である．

最後に，可逆な保測変換 T_1, T_2 をもつ力学系 $(X_1, \mathcal{F}_1, \mu_1, T_1)$ と $(X_2, \mathcal{F}_2, \mu_2, T_2)$ の同型問題に関する von Neumann の結果を述べよう．各 T_i に対して，$(U_i f)(x)=f(T_i x)$ $(\forall x \in X_i,\ \forall f \in L^2(X_i))$ とおくと，U_i は Hilbert 空間 $L^2(X_i)$ 上のユニタリー作用素となる．ここで，U_i が離散スペクトルをもつとき，T_i は**離散スペクトル**をもつという．さらに，U_1 と U_2 がユニタリー同値（すなわち，ユニタリー作用素 $V: L^2(X_1) \to L^2(X_2)$ が存在して，$U_1=V^* U_2 V$）であるとき，T_1 と T_2 を**スペクトル同型**といい，$T_1 \simeq T_2$ で表す．このとき，"$T_1 \cong T_2 \Rightarrow T_1 \simeq T_2$" は容易にわかるが，逆はどうかという問題が von Neumann, Halmos 以来の懸案であった．彼の結果は

〈**定理 2.14**〉 エルゴード的な T_1, T_2 に対して，T_1 または T_2 が離散スペクトルをもつとき，$T_1 \cong T_2$ であることと，$T_1 \simeq T_2$ は同値．

註 変換 T_i がエルゴード的であるとは，T_i に関して不変な可測集合はつねに測度 0 か 1 であるときをいう．この定理の証明は von Neumann の論文集 [V.1, Vol. II] または [H.3] を参照．上述の同型問題に関しては，各 T_i のエントロピー $S(T_i)$ の関連で，さらにさまざまな結果が得られている．また，この節の議論と関連して Kolmogorov 変換も数学上重要な概念であるが，ここでは割愛した．

第3章　相対エントロピー

この章では，２つの確率測度の間の不確定さを表す相対エントロピーを導入し，その主要な性質を論じる．また，統計学において重要な概念である十分性と相対エントロピーの関連について概説する．

3.1　相対エントロピーの導入

可測定間 $(\Omega, \mathcal{F})$ 上の２つの確率測度（以下の議論の殆どは有限測度でも成立）μ, ν に関して

$$S_{\mathcal{G}}(\mu|\nu) = \sup\left\{\sum_{A \in \tilde{A}} \mu(A) \log \frac{\mu(A)}{\nu(A)};\ \tilde{A} \in \mathcal{P}(\mathcal{G})\right\} \tag{3.1}$$

を，$\mathcal{F}$ の σ-部分集合体 $\mathcal{G}$ **に関する μ と ν の相対エントロピー**という．とくに $S_{\mathcal{F}}(\mu|\nu)$ を単に $S(\mu|\nu)$ とかき，**μ と ν の相対エントロピー**という．上の定義から，次は明らかである．

1°　**正値性**：$S(\mu|\nu) \geqq 0$，$S(\mu|\nu) = 0 \Longleftrightarrow \mu = \nu$.

2°　**単調性**：$\mathcal{G}_1 \subset \mathcal{G}_2$ のとき，$S_{\mathcal{G}_1}(\mu|\nu) \leqq S_{\mathcal{G}_2}(\mu|\nu) \leqq S(\mu|\nu)$.

〈定理 3.1〉 Gelfand-Kolmogorov-Yaglom の定理：可測空間 $(X, \mathcal{F})$ 上の２つの確率測度 μ, ν に関して

（１）　$\mu \ll \nu$ ならば

$$S(\mu|\nu) = \int_\Omega \left(\frac{d\mu}{d\nu} \log \frac{d\mu}{d\nu}\right) d\nu = \int_\Omega \log \frac{d\mu}{d\nu}\, d\mu. \tag{3.2}$$

（２）　$\mu \not\ll \nu$ ならば，$S(\mu|\nu) = +\infty$.

証明　（１）：函数 $t \longmapsto t \log t$ $(t \geqq 0)$ の凸性より，$\nu(E) \neq 0$ である任意の $E \in \mathcal{F}$ と可測函数 $f \geqq 0$ に対して，

$$\left(\frac{1}{\nu(E)}\int_E f\,d\nu\right)\log\left(\frac{1}{\nu(E)}\int_E f\,d\nu\right)\leqq\frac{1}{\nu(E)}\int_E f\log f\,d\nu.$$

ここで，$f=d\mu/d\nu$ とおくと，

$$\mu(E)\log\frac{\mu(E)}{\nu(E)}\leqq\int_E\left(\frac{d\mu}{d\nu}\log\frac{d\mu}{d\nu}\right)d\nu.$$

任意の有限分割 $\tilde{A}\in\mathcal{P}(\mathcal{F})$ に対して，

$$\sum_{E\in\tilde{A}}\mu(E)\log\frac{\mu(E)}{\nu(E)}\leqq\int_\Omega\left(\frac{d\mu}{d\nu}\log\frac{d\mu}{d\nu}\right)d\nu.$$

ゆえに，sup をとると，

$$S(\mu|\nu)\leqq\int_\Omega\left(\frac{d\mu}{d\nu}\log\frac{d\mu}{d\nu}\right)d\nu.$$

したがって，逆向きの不等式を証明する．今，

$$E_{n,k}=\left\{\omega\in\Omega;\ \frac{k}{n}\leqq\frac{d\mu}{d\nu}(\omega)<\frac{k+1}{n}\right\}\quad(k=0,1,\cdots,n^2-1),$$

$$E_{n,n^2}=\left\{\omega\in\Omega;\ \frac{d\mu}{d\nu}(\omega)\geqq n\right\}$$

とおくと，$\{E_{n,k};\ k=0,1,\cdots,n^2\}$ は Ω の有限分割となる．$E_{n,k}$ の定義より

$$\frac{k}{n}\nu(E_{n,k})\leqq\int 1_{E_{n,k}}\frac{d\mu}{d\nu}d\nu\leqq\frac{k+1}{n}\nu(E_{n,k}).$$

ゆえに，

$$\frac{k}{n}\leqq\frac{\mu(E_{n,k})}{\nu(E_{n,k})}<\frac{k+1}{n}.$$

ここで，

$$K_1=\left\{k;\ \frac{k+1}{n}\leqq e^{-1}\right\},$$

$$K_2=\left\{k;\ \frac{k}{n}\geqq e^{-1}\right\}$$

とおくと，函数 $t\log t$ の性質より

$$k\in K_1 \Rightarrow \frac{k+1}{n}\log\frac{k+1}{n}\leqq\frac{\mu(E_{n,k})}{\nu(E_{n,k})}\log\frac{\mu(E_{n,k})}{\nu(E_{n,k})},$$

$$k\in K_2 \Rightarrow \frac{k}{n}\log\frac{k}{n}\leqq\frac{\mu(E_{n,k})}{\nu(E_{n,k})}\log\frac{\mu(E_{n,k})}{\nu(E_{n,k})}.$$

さて，函数

$$f_n(\omega)=\sum_{k=0}^{n^2-1}\frac{k}{n}1_{E_{n,k}}(\omega), \qquad g_n(\omega)=\sum_{k=0}^{n^2-1}\frac{k+1}{n}1_{E_{n,k}}(\omega)$$

に対して

$$h_n(\omega)=\begin{cases} f_n(\omega) & (\omega\in \bigcup_{k\in K_2} E_{n,k}), \\ g_n(\omega) & (\omega\in \bigcup_{k\in K_1} E_{n,k}) \end{cases}$$

とおくと，

$$\sum_{k=0}^{n^2}\mu(E_{n,k})\log\frac{\mu(E_{n,k})}{\nu(E_{n,k})}\geqq\sum_{k\in K_2}\int_{E_{n,k}}f_n\log f_n\,d\nu+\sum_{k\in K_1}\int_{E_{n,k}}g_n\log g_n\,d\nu$$
$$=\int_\Omega h_n\log h_n\,d\nu.$$

したがって，

$$S(\mu|\nu)\geqq\int_\Omega h_n\log h_n\,d\nu.$$

ところで，$n\to\infty$ のとき，任意の $\omega\in\Omega$ に対して，$f=\dfrac{d\mu}{d\nu}$ とすると

$$0\leqq g_n(\omega)-f_n(\omega)\leqq\frac{1}{n}\to 0,$$

$$0\leqq f(\omega)-f_n(\omega)\leqq g_n(\omega)-f_n(\omega)\leqq\frac{1}{n}\to 0,$$

$$0\leqq g_n(\omega)-f(\omega)\leqq g_n(\omega)-f_n(\omega)\leqq\frac{1}{n}\to 0$$

であるから，

$$0\leqq|f-h_n|\leqq|f-f_n|1_{\bigcup_{k\in K_2}E_{n,k}}+|f-g_n|1_{\bigcup_{k\in K_1}E_{n,k}}$$

より

$$\int_\Omega f\log f\,d\nu=\lim_{n\to\infty}\int_\Omega h_n\log h_n\,d\nu\leqq S(\mu|\nu).$$

（2）は自明であろう．▌

この定理の証明は，最初 Kallianpour ［K.3］ で与えられた．

さて，確率測度 μ,ν が密度函数をもつ場合において $S(\mu|\nu)$ を論じよう．確率空間 $(\Omega,\mathcal{F},\mu)$ 上の実数値確率変数 ξ に対して，実数直線 $\boldsymbol{R}=(-\infty,\infty)$

上の確率測度（ξ の分布）$\mu_\xi(A)=\mu(\xi^{-1}(A))$ $(\forall A\in\mathcal{B}(\mathcal{B}_R))$ が定まる：これが $\mu_\xi(dt)\ll dt(=\text{Lebesgne}$ 測度)，つまり ξ の分布が絶対連続のとき，ξ を絶対連続な確率変数といい，その Radon-Nikodym 微分 $f(t)=\mu_\xi(dt)/dt$ を ξ の**密度函数**という．2つの絶対連続な確率変数 ξ,ζ に対する密度函数を f,g とするとき，

$$I(\xi|\zeta)=\int_R(f(t)\log f(t)-f(t)\log g(t))dt$$

を ξ と ζ の **Kullback-Leibler の情報量**とよぶ．これについて，まず次の補題（基本不等式）を得る．

〈**補題** 3.2〉 下記の両辺が有限の場合，不等式

$$-\int_R f(t)\log f(t)dt\leq-\int_R f(t)\log g(t)dt \tag{3.3}$$

が成立する．等号が成立するための必要十分条件は，$f(t)=g(t)$ a.e. である．

証明 一般に $\log\lambda\leq\lambda-1$ $(\forall\lambda>0)$ であり，等号は $\lambda=1$ の場合に成立する．ゆえに，

$$f(t)(\log g(t)-\log f(t))=f(t)\log\frac{g(t)}{f(t)}\leqslant g(t)-f(t)\ \text{a.e.}.$$

ここで，$a/0=\infty$，$0\cdot\infty=0$ と規約する．ゆえに，

$$\int f(t)\log g(t)dt-\int f(t)\log f(t)dt\leqslant\int g(t)dt-\int f(t)dt=1-1=0.$$

さらに，等号が成立するのは，$f(t)=g(t)$ a.e. の場合に限られる．▮

Kullback-Leibler の情報量 $I(\xi|\zeta)$ と $\boldsymbol{R}$ 上の確率測度 μ_ξ,μ_ζ 間の相対エントロピーの間の関係は，いずれか一方が有限ならば，両者はつねに一致する，すなわち，

$$I(\xi|\zeta)=S(\mu_\xi|\mu_\zeta).$$

実際，$S(\mu_\xi|\mu_\zeta)<\infty$ ならば，$\mu_\xi\ll\mu_\zeta$ である．このことより，

$$\begin{aligned}I(\xi|\zeta)&=\int_R(f(t)\log f(t)-f(t)\log g(t))dt\\&=\int_R f(t)\left(\log\frac{f(t)}{g(t)}\right)dt\end{aligned}$$

$$=\int_R \frac{d\mu_\xi}{dt}\log\left(\frac{d\mu_\xi}{dt}\Big/\frac{d\mu_\zeta}{dt}\right)dt$$
$$=\int_R \left(\frac{d\mu_\xi}{d\mu_\zeta}\log\frac{d\mu_\xi}{d\mu_\zeta}\right)d\mu_\zeta=S(\mu_\xi|\mu_\zeta).$$

更に，$I(\xi|\zeta)<\infty$ のとき，同じ計算式を吟味すれば，容易に $\mu_\xi\ll\mu_\zeta$ がいえ，上記の等式がすべて成立する．

3.2　相対エントロピーの諸性質

この節では $S(\mu|\nu)$ の主要な性質を論じる．以下，μ,ν 等は可測空間 $(\Omega,\mathcal{F})$ 上の確率測度とする．

〈定理 3.3〉 $\mathcal{F}$ の σ-部分集合体の列 $\mathcal{G}_n\uparrow\mathcal{G}$ に対して，

$$S_{\mathcal{G}_n}(\mu|\nu)\uparrow S_{\mathcal{G}}(\mu|\nu).$$

証明　相対エントロピーの単調性 2° より，

$$S_{\mathcal{G}_n}(\mu|\nu)\leqslant S_{\mathcal{G}_{n+1}}(\mu|\nu)\leqslant S_{\mathcal{G}}(\mu|\nu)$$

であるから，収束

$$S_{\mathcal{G}_n}(\mu|\nu)\longrightarrow S_{\mathcal{G}}(\mu|\nu)$$

を示せばよい．まず，$\mathcal{F}$ 上で $\mu\ll\nu$ のときを考える．

なお，この定理の証明では $\mathcal{G}=\mathcal{F}$ としても一般性を失わないので，以下，これを仮定する．$\mu_n=\mu\restriction\mathcal{G}_n$ （$\mathcal{G}_n$ への制限），$\nu_n=\nu\restriction\mathcal{G}_n$ とおくと，$\mu\ll\nu$ より $\mu_n\ll\nu_n$ となる．よって，$f_n=d\mu_n/d\nu_n$，$f=d\mu/d\nu$ とおくと，〈定理 3.1〉より，$S_{\mathcal{G}_n}(\mu|\nu)=\int f_n\log f_n\,d\nu$，$S_{\mathcal{G}}(\mu|\nu)=\int f\log f\,d\nu$ となり，$\{f_n,f\}$ は martingale で $f_n\to f$ （ν-a.e.）であるから，$f_n\log f_n\to f\log f$ （ν-a.e.）となる．したがって，Fatou の補題から

$$S_{\mathcal{G}}(\mu|\nu)=\int f\log f\,d\nu\leqq\liminf_{n\to\infty}\int f_n\log f_n\,d\nu$$
$$=\liminf_{n\to\infty}S_{\mathcal{G}_n}(\mu|\nu)\leqq S_{\mathcal{G}}(\mu|\nu).$$

次に，$\mu\not\ll\nu$ の場合を考える．このとき，$\nu(A)=0$ かつ $\mu(A)>0$ である $A\in\mathcal{G}$ が存在する．そこで，$\mu(A)>2\varepsilon$ を満たす任意の $\varepsilon>0$ に対して，$\mathcal{G}_n\uparrow\mathcal{G}$

であるから，ある番号 $n\in N$ と $B\in\mathcal{G}_n$ が存在して，$\mu(A\triangle B)<\varepsilon$，$\nu(A\triangle B)<\varepsilon$（$\triangle$は対称差を表す）．今，$\tilde{B}=\{B, B^c\}\in\mathcal{P}(\mathcal{G}_n)$ とおくと，

$$\lim_{n\to\infty} S_{\mathcal{G}_n}(\mu|\nu)\geqq S_{\mathcal{G}_n}(\mu|\nu)\geqq S_{\tilde{B}}(\mu|\nu)$$

$$=\mu(B)\log\frac{\mu(B)}{\nu(B)}+\mu(B^c)\log\frac{\mu(B^c)}{\nu(B^c)}. \tag{3.4}$$

ところで，

$$\mu(B)>\mu(A)-\varepsilon>\frac{1}{2}\mu(A), \qquad \nu(B)<\nu(A)+\varepsilon=\varepsilon$$

であるから，

$$\mu(B)\log(\mu(B)/\nu(B))\geqq\frac{1}{2}\mu(A)\log(\mu(B)/\varepsilon)$$

および

$$\mu(B^c)\log(\mu(B^c)/\nu(B^c))\geqq\mu(B^c)(1-\nu(B^c)/\mu(B^c))=\mu(B^c)-\nu(B^c)$$

となる．したがって，(3.4) は，$\varepsilon\to 0$ をとると

$$\lim_{n\to\infty} S_{\mathcal{G}_n}(\mu|\nu)\geqq\frac{1}{2}\mu(A)\log(\mu(B)/\varepsilon)+\mu(B^c)-\nu(B^c)\to\infty.$$

ゆえに，

$$\lim_{n\to\infty} S_{\mathcal{G}_n}(\mu|\nu)=\infty=S_{\mathcal{G}}(\mu|\nu). \quad \blacksquare$$

〈定理 3.4〉 同時凸性：任意の $\alpha\in[0,1]$ に対して，

$$S(\alpha\mu_1+(1-\alpha)\mu_2|\alpha\nu_1+(1-\alpha)\nu_2)\leqq\alpha S(\mu_1|\nu_1)+(1-\alpha)S(\mu_2|\nu_2).$$

証明 $\alpha=0,1$ のときは自明だから，$0<\alpha<1$ とする．$\mu_1\not\ll\nu_1$，あるいは $\mu_2\not\ll\nu_2$ であるときは，それぞれ $S(\mu_1|\nu_1)=\infty$，$S(\mu_2|\nu_2)=\infty$ となり，不等式は当然成立する．$\mu_1\ll\nu_1$，$\mu_2\ll\nu_2$ の場合を証明する．定理の不等式の証明は，任意の $\tilde{A}\in\mathcal{P}(\mathcal{F})$ に対して，

$$\sum_{A\in\tilde{A}}\{\alpha\mu_1(A)+(1-\alpha)\mu_2(A)\}\log\frac{\alpha\mu_1(A)+(1-\alpha)\mu_2(A)}{\alpha\nu_1(A)+(1-\alpha)\nu_2(A)}$$

$$\leqq\alpha\sum_{A\in\tilde{A}}\mu_1(A)\log\frac{\mu_1(A)}{\nu_1(A)}+(1-\alpha)\sum_{A\in\tilde{A}}\mu_2(A)\log\frac{\mu_2(A)}{\nu_2(A)} \tag{3.5}$$

を示すことと同値．このためには，ある定数 $\mu_1, \mu_2, \nu_1, \nu_2$ に対して，函数

$$f(x)=\{x\mu_1+(1-x)\mu_2\}\log\frac{x\mu_1+(1-x)\mu_2}{x\nu_1+(1-x)\nu_2}$$

が凸であることを示せばよい．実際，$f(x)$ が凸函数であれば，

$$f(\alpha\cdot 1+(1-\alpha)\cdot 0)\leqq \alpha f(1)+(1-\alpha)f(0) \quad (\forall \alpha\in(0,1))$$

となり，(3.5) が示される．ところで，上の函数 $f(x)$ は $(0,1)$ で微分可能，かつ $f''(x)\geqq 0$ であるから，$f(x)$ は $[0,1]$ で凸函数である．▮

可測空間 $(\Omega, \mathcal{F})$ 上の符号付測度 λ の全変動ノルムを $\|\lambda\|$ とする．このノルムに関して，測度の収束を定めると，$S(\mu|\nu)$ の連続性が考えられる．

〈定理 3.5〉　下半連続性：$\|\mu_n-\mu\|\to 0, \|\nu_n-\nu\|\to 0$ のとき，

$$S(\mu|\nu)\leqq \liminf_{n\to\infty} S(\mu_n|\nu_n).$$

証明　任意の $\tilde{A}\in\mathcal{P}(\mathcal{F})$ に対して

$$\liminf S_{\tilde{A}}(\mu_n|\nu_n)\geqq S_{\tilde{A}}(\mu|\nu)$$

を示せばよい．条件より，$\mu_n(A)\to\mu(A)$，$\nu_n(A)\to\nu(A)$　$(\forall A\in\tilde{A})$ であるから，

$$\nu(A)>0 \Rightarrow \lim_{n\to\infty}\mu_n(A)\log(\mu_n(A)/\nu_n(A))=\mu(A)\log(\mu(A)/\nu(A)),$$

$$\nu(A)=0,\ \mu(A)>0 \Rightarrow \lim_{n\to\infty}\mu_n(A)\log(\mu_n(A)/\nu_n(A))=\infty$$

$$=\mu(A)\log(\mu(A)/\nu(A)),$$

$$\nu(A)=\mu(A)=0 \Rightarrow \mu_n(A)\log(\mu_n(A)/\nu_n(A))\geqq \mu_n(A)-\nu_n(A)\to 0$$

$$\Rightarrow \liminf_{n\to\infty}\mu_n(A)\log(\mu_n(A)/\nu_n(A))\geqq 0=\mu(A)\log\frac{\mu(A)}{\nu(A)}.$$

ゆえに，$S_{\tilde{A}}(\mu_n|\nu_n)$ の定義より

$$\liminf_{n\to\infty} S_{\tilde{A}}(\mu_n|\nu_n)\geqq S_{\tilde{A}}(\mu|\nu).\ ▮$$

3.3　十分性と相対エントロピー

以下，導入する十分性の概念は，統計的決定や推定などの問題において重要な役割を演じている（たとえば，文献 [N.1] 参照）．この節での議論のほとん

どは Halmos–Savage [H. 1] および Kullback–Leibler [K. 15] によっている.

可測空間 $(\Omega, \mathcal{F})$ 上の確率測度の全体を $\boldsymbol{P}(\Omega)$ で表し，$\boldsymbol{P}(\Omega)$ の部分集合を $\boldsymbol{P}, \boldsymbol{Q}$ 等で表す. 2つの $\boldsymbol{P}(\Omega)$ の部分集合 $\boldsymbol{P}_1, \boldsymbol{P}_2$ に関して

（1） $\boldsymbol{P}_1$ が $\boldsymbol{P}_2$ を**押さえる**（$\boldsymbol{P}_1 \gg \boldsymbol{P}_2$ または $\boldsymbol{P}_2 \ll \boldsymbol{P}_1$ とかく）とは，$A \in \mathcal{F}$ がすべての $\mu \in \boldsymbol{P}_1$ に対して $\mu(A)=0$ ならば，すべての $\nu \in \boldsymbol{P}_2$ に対しても $\nu(A)=0$ となることである（この場合，$\boldsymbol{P}_2$ は $\boldsymbol{P}_1$ に**よって押さえられているともいう**）.

（2） $\boldsymbol{P}_1$ と $\boldsymbol{P}_2$ とが**同等**（$\boldsymbol{P}_1 \sim \boldsymbol{P}_2$ とかく）とは，$\boldsymbol{P}_1 \gg \boldsymbol{P}_2$ かつ $\boldsymbol{P}_2 \gg \boldsymbol{P}_1$ となることである.

（3） $\boldsymbol{P}_1$ が**押さえられた集合**であるとは，$\lambda \in \boldsymbol{P}(\Omega)$ が存在して，$\boldsymbol{P}_1 \ll \{\lambda\}$ となることである. 以下，このような集合は $\boldsymbol{D}$ で表すことにする. すなわち，この節ではとくに断らない限り，$\boldsymbol{D}$ は押さえられた集合である.

〈補題 3.6〉 $\boldsymbol{D}$ はつねに $\boldsymbol{P}(\Omega)$ の高々可算な部分集合と同等である.

証明 $\boldsymbol{D} \ll \{\lambda\}$，$f_\mu = d\mu/d\lambda$ $(\forall \mu \in \boldsymbol{D})$，$K_\mu = \{\omega \in \Omega;\, f_\mu(\omega) > 0\}$ とおく. ある $\mu \in \boldsymbol{D}$ に対して，$K \subset K_\mu$，$\mu(K) > 0$ を満たす Ω の部分集合 K を考える. 今，この集合 K を核とよぶことにして，たがいに素な核の和集合を鎖とよぶことにする. $\lambda(\Omega) < \infty$ であるから，鎖はたがいに素な核の高々可算個の和集合で表される. また，たがいに素な鎖の和集合も鎖である. したがって，鎖の可算個（たがいに素でなくてもよい）の和集合は鎖となる. ところで，$\lambda(C_k) \uparrow \lambda(C)$ である鎖の列 $\{C_k\}$ は，$C = \bigcup_{k=1}^{\infty} C_k$ と表せ，C も鎖となる. このとき，鎖の定義より核の列 $\{K_n\}$ が存在して，$C = \bigcup_{n=1}^{\infty} K_n$ とかけ，また，核の定義より，各 K_n に対して，$\mu_n \in \boldsymbol{D}$ が存在して，$K_n \subset K_{\mu_n}$，$\mu_n(K_n) > 0$ となる. ここで，$\widetilde{\boldsymbol{D}} = \{\mu_1, \mu_2, \cdots\}$ とおくと，$\widetilde{\boldsymbol{D}} \subset \boldsymbol{D}$ であるから，$\widetilde{\boldsymbol{D}} \ll \boldsymbol{D}$ である. よって，逆に $\boldsymbol{D} \ll \widetilde{\boldsymbol{D}}$ を示せばよい. 今，$A \in \mathcal{F}$ が $\mu_n(A) = 0$ $(\forall n)$ を満たすとする. このとき，任意の $\mu \in \boldsymbol{D}$ に対して，K_μ の定義より，$\mu(A - K_\mu) = 0$ となるから $A \subset K_\mu$ と仮定してよい. このことから，鎖 C に対して，容易に $\mu(A - C) = 0$ かつ $\mu(A \cap C) = 0$ が示せる. すなわち，$\mu(A) = 0$. ゆえに $\boldsymbol{D} \ll \widetilde{\boldsymbol{D}}$. ∎

さて，十分性の概念を導入しよう．$\mathcal{G}$ を $\mathcal{F}$ の σ-部分集合体とし，$\boldsymbol{P}$ を $\boldsymbol{P}(\Omega)$ の部分集合とするとき，$\mathcal{G}$ が $\boldsymbol{P}$ に関して**十分** (sufficient) であるとは，任意の $A\in\mathcal{F}$ に対して，$\mathcal{G}$-可測函数 h が存在して，

$$h=E_{\mu}(1_A|\mathcal{G}) \quad \mu\text{-a.e.}, \qquad \forall\mu\in\boldsymbol{P} \tag{3.6}$$

となることをいう．すなわち，a.e. の意味で，各 $\mu\in\boldsymbol{P}$ の $\mathcal{G}$ に関する条件付確率 ($P_{\mu}(A|\mathcal{G})=E_{\mu}(1_A|\mathcal{G})$ より）が等しい場合を，$\mathcal{G}$ が $\boldsymbol{P}$ に関して十分であるというのである．

なお，以下の議論では $\mathcal{G}$ に関する可測函数の全体を $\mathfrak{M}_{\mathcal{G}}$ とかくことにする．

〈定理 3.7〉 $\mathcal{G}$ が $\boldsymbol{D}$ に関して十分であることと $\boldsymbol{D}\sim\{\lambda\}$ かつ $d\mu/d\lambda\in\mathfrak{M}_{\mathcal{G}}$ (λ-a.e., $\forall\mu\in\boldsymbol{D}$) を満たす $\lambda\in\boldsymbol{P}(\Omega)$ が存在することとは同値．

証明 必要性の証明：〈補題 3.6〉より，$\boldsymbol{D}=\{\mu_1,\mu_2,\cdots\}$ としてよい．また，

$$\lambda(A)=\sum_{k=1}^{\infty}\frac{1}{2^k}\mu_k(A), \qquad \forall A\in\mathcal{F}$$

とおくと，明らかに，$\lambda\in\boldsymbol{P}(\Omega)$ かつ $\boldsymbol{D}\sim\{\lambda\}$．さらに，任意の $A\in\mathcal{F}, B\in\mathcal{G}$ に対して，$\mathcal{G}$ が $\boldsymbol{D}$ に関して十分であるから，$h\in\mathfrak{M}_{\mathcal{G}}$ が存在して

$$\lambda(A\cap B)=\sum_{k=1}^{\infty}\frac{1}{2^k}\int_B E_{\mu_k}(1_A|\mathcal{G})\,d\mu_k=\int_B h\,d\lambda$$

となる．よって $E_{\lambda}(1_A|\mathcal{G})=h$ λ-a.e. でもある．このことから，任意の $\mu\in\boldsymbol{D}$ に対して，$g=d\mu/d\lambda$ とおくと，

$$\int_A g\,d\lambda=\int_A E_{\lambda}(g|\mathcal{G})\,d\lambda, \qquad \forall A\in\mathcal{F}$$

が簡単にわかる．したがって $g=E_{\lambda}(g|\mathcal{G})$ (λ-a.e.)，すなわち $g\in\mathfrak{M}_{\mathcal{G}}$ (λ-a.e.) である．

逆向き，十分性の証明：$\mu\in\boldsymbol{D}$ と $A\in\mathcal{F}$ および $B\in\mathcal{G}$ に対して，$g\equiv d\mu/d\lambda\in\mathfrak{M}_{\mathcal{G}}$ (λ-a.e.) であるから，

$$\begin{aligned}\int_B E_{\lambda}(1_A|\mathcal{G})\,d\mu&=\int_B E_{\lambda}(1_A|\mathcal{G})\,g\,d\lambda\\&=\int_B 1_A\cdot g\,d\lambda=\int_B E_{\mu}(1_A|\mathcal{G})\,d\mu.\end{aligned}$$

$B \in \mathcal{G}$ は任意であるから，

$$E_\lambda(1_A|\mathcal{G}) = E_\mu(1_A|\mathcal{G}) \ \mu\text{-a.e.}, \qquad \forall \mu \in \boldsymbol{D}.$$

ゆえに，$\mathcal{G}$ は $\boldsymbol{D}$ に関して十分である．▌

さて，Ω 上の確率測度の全体 $\boldsymbol{P}(\Omega)$ の部分集合 $\boldsymbol{P}$ が**一様**であるとは，$\boldsymbol{P}$ の任意の元 μ, ν がたがいに絶対連続である場合をいう．この一様な部分集合 $\boldsymbol{P}$ は押さえられた確率測度の集合 $\boldsymbol{D}$ に含まれる．さらに，上記の定理より次の系が容易に求まる．

〈系 3.8〉 $\boldsymbol{P}$ が一様のとき，$\mathcal{G}$ が $\boldsymbol{P}$ に関して十分であることと，$d\nu/d\mu \in \mathfrak{M}_\mathcal{G}$ (μ-a.e., $\forall \mu, \nu \in \boldsymbol{P}$) であることとは同値である．

2つの測度 $\{\mu, \nu\}$ の $\mathcal{G}$ に関する十分性，いわゆる**対十分性**は次のように定義される．

(4) $\mathcal{G}$ が $\boldsymbol{P}$ に関して**対十分**とは，任意の対 $\mu, \nu \in \boldsymbol{P}$ に関して，$\mathcal{G}$ が十分となることである．

〈補題 3.9〉 $\mathcal{G}$ が $\boldsymbol{P}$ に関して対十分であることと，$\boldsymbol{P}$ の任意の対 μ, ν において $d\mu/d(\mu+\nu) \in \mathfrak{M}_\mathcal{G}$ ($\mu+\nu$-a.e.) であることとは同値である．

証明 $\mathcal{G}$ が $\boldsymbol{P}$ に関して対十分であるとき，定義より，任意の $\mu, \nu \in \boldsymbol{P}$ に対して，$\mathcal{G}$ は $\{\mu, \nu\}$ に関して十分．よって，$\lambda = \dfrac{1}{2}(\mu+\nu)$ とおくと，〈定理 3.7〉の証明と同じ議論により，$\dfrac{d\mu}{d\lambda}$, $\dfrac{d\nu}{d\lambda} \in \mathfrak{M}_\mathcal{G}$．したがって，$\dfrac{d\mu}{d(\mu+\nu)} \in \mathfrak{M}_\mathcal{G}$ ($\mu+\nu$-a.e.)．逆は，〈定理 3.7〉より明らか．▌

〈定理 3.10〉 $\mathcal{G}$ が $\boldsymbol{D}$ に関して十分であることと，$\mathcal{G}$ が $\boldsymbol{D}$ に関して対十分であることとは同値である．

証明 $\mathcal{G}$ が $\boldsymbol{D}$ に関して十分ならば，定義より，$\mathcal{G}$ は $\boldsymbol{D}$ に関して対十分である．逆に，$\mathcal{G}$ が $\boldsymbol{D}$ に関して対十分であるとする．〈補題 3.6〉より $\boldsymbol{D} = \{\mu_1, \mu_2, \cdots\}$ としてよい．ここで，

$$\lambda(A)=\sum_{k=1}^{\infty}\frac{1}{2^k}\mu_k(A)\quad(\forall A\in\mathcal{F})$$

とおくと，任意の$\mu\in\boldsymbol{D}$に対して，〈補題 3.9〉より$d\mu/d(\mu+\lambda)\in\mathfrak{M}_{\mathcal{G}}$ ($\mu+\lambda$-a.e.) がわかる．したがって，

$$\frac{d\mu}{d\lambda}=\frac{d\mu}{d(\mu+\lambda)}\left(1-\frac{d\mu}{d(\mu+\lambda)}\right)^{-1}$$

となるから，$d\mu/d\lambda\in\mathfrak{M}_{\mathcal{G}}$．よって，〈定理 3.7〉より，$\mathcal{G}$は$\boldsymbol{D}$に関して十分である．▮

さて，ここで，相対エントロピーに関する議論にもどろう．Kullback と Leibler は次の定理を証明した．

〈定理 3.11〉　$\mu,\nu\in\boldsymbol{P}(\Omega)$ において，次が成立する．

(1)　$\mathcal{G}$が$\{\mu,\nu\}$に関して十分ならば，$S_{\mathcal{G}}(\mu|\nu)=S(\mu|\nu)$．

(2)　$S_{\mathcal{G}}(\mu|\nu)=S(\mu|\nu)<+\infty$ ならば，$\mathcal{G}$は$\{\mu,\nu\}$に関して十分である．

証明　$\mu\ll\nu$ の場合，命題を証明すればよい．μ_0,ν_0 を各々μとνの$\mathcal{G}$への制限とする．このとき，

$$S(\mu|\nu)-S_{\mathcal{G}}(\mu|\nu)=\int_{\Omega}\log\left(\frac{d\mu}{d\nu}\Big/\frac{d\mu_0}{d\nu_0}\right)d\mu$$

$$=\int_{\Omega}\left(\frac{d\mu}{d\nu}\Big/\frac{d\mu_0}{d\nu_0}\right)\log\left(\frac{d\mu}{d\nu}\Big/\frac{d\mu_0}{d\nu_0}\right)\frac{d\mu_0}{d\nu_0}d\nu$$

となる．ここで，

$$f=\left(\frac{d\mu}{d\nu}\Big/\frac{d\mu_0}{d\nu_0}\right),\qquad d\xi=\left(\frac{d\mu_0}{d\nu_0}\right)d\nu$$

とおくと，$\int_{\Omega}f\,d\xi=1$ であるから，$\int_{\Omega}f\log f\,d\xi\geqq 0$ で，等号は $f=1$ (ξ-a.e.) のとき成立することが容易にわかる．したがって，$S_{\mathcal{G}}(\mu|\nu)=S(\mu|\nu)$ であることと，$f=1$ すなわち $d\mu/d\nu=d\mu_0/d\nu_0$ が成立することは同値である．さらに，$\lambda=\frac{1}{2}(\mu+\nu)$ とおくと，$\{\mu,\nu\}\sim\lambda$ である．したがって，〈定理 3.7〉より (1),(2) が成立する．▮

Csiszar [C.1] は2つの確率測度μ,νの間の距離が，それらの相対エントロピーによって押さえられることを示した．

〈定理 3.12〉 任意の $\mu, \nu \in \boldsymbol{P}(\Omega)$ に関して，

$$\|\mu-\nu\| \leqq \sqrt{2S(\mu|\nu)}.$$

証明 任意の $A \in \mathcal{F}$ に対して，Hahn 分解によって，$F \in \mathcal{F}$ が存在して，$(\mu-\nu)^+(A)=(\mu-\nu)(A\cap F)$ かつ $(\mu-\nu)^-(A)=(\nu-\mu)(A\cap F^c)$ と表せる．このとき，

$$\|\mu-\nu\|=(\mu-\nu)^+(\Omega)-(\mu-\nu)^-(\Omega)=2(\mu(F)-\nu(F)).$$

さて，$\tilde{A}=(F, F^c) \in \mathcal{P}(\mathcal{F})$ とおくと，

$$S_{\tilde{A}}(\mu|\nu)=\mu(F)\log\frac{\mu(F)}{\nu(F)}+(1-\mu(F))\log\frac{1-\mu(F)}{1-\nu(F)}.$$

ところで，$0\leqq x\leqq y\leqq 1$ なる任意の x, y に対して，

$$\sqrt{2\left\{y\log\frac{y}{x}+(1-y)\log\frac{1-y}{1-x}\right\}}\geqq 2(y-x)$$

が成立する．これは簡単な微分の計算で示せる．ここで，$x=\nu(F)$，$y=\mu(F)$ とおくと，

$$\sqrt{2S(\mu|\nu)}\geqq\sqrt{2S_{\tilde{A}}(\mu|\nu)}\geqq 2(\mu(F)-\nu(F))=\|\mu-\nu\|$$

となる．▮

上の定理において，$\sqrt{\ }$ の中の数 2 は，この不等式が成立するための最小数であることに注意しよう．この証明も初等的である．

なお，$(X, \mathcal{F})$，$(Y, \mathcal{G})$ を 2 つの可測空間とし，T を X から Y への可測変換とするとき，Halmos–Savage は十分性を次のように定義している [H.1]．今，$\boldsymbol{P}$ を $\mathcal{F}$ 上の確率測度のある集合とするとき，変換 T が $\boldsymbol{P}$ に関して**十分**であるとは，任意の $A \in \mathcal{F}$ と $\mu \in \boldsymbol{P}$ に対して，$h \in \mathfrak{M}_{\mathcal{G}}$ が存在して，

$$h=E_{\mu}(1_A|\mathcal{G}) \quad \mu\circ T^{-1}\text{-a.e.}$$

となる場合をいう．ここでは，上の $E_{\mu}(\cdot|\mathcal{G})$ は次のように定められる $\mathcal{G}$-可測函数である．

$$\int_{T^{-1}(G)} f(x)d\mu(x)=\int_G E_{\mu}(f|\mathcal{G})(y)d\mu\circ T^{-1}(y), \qquad \forall G\in\mathcal{G},\ \forall f\in\mathfrak{M}_{\mathcal{F}}.$$

変換 T が X から X への可測写像で，$\mathcal{G}=T^{-1}(\mathcal{F})$ である場合が，我々が今まで論じてきたものである．T を陽に含む場合は，たとえば〈定理 3.7〉は次

のようにかき表せる.

「T が $\boldsymbol{D}$ に関して十分であることと, $\boldsymbol{D}\sim\{\lambda\}$ かつ

$$\frac{d\mu}{d\lambda}\in\mathfrak{M}_{T^{-1}(\mathcal{G})}\quad(\lambda\text{-a. e.},\ \ \forall\mu\in\boldsymbol{D})$$

を満たす $\lambda\in\boldsymbol{P}(X)$ が存在することとは同値である」

このように, この節でのすべての定理は, T に関するそれにかき換えることができる.

ところで, $\mathcal{G}$ が $\boldsymbol{P}$ に関して十分であることの直観的意味は, $\boldsymbol{P}$ に属する任意の測度 μ,ν の違いが $\mathcal{G}$ によって完全に決定されるということであり, 言い換えれば, $\mathcal{G}$-可測函数(を測定すること)によって $\boldsymbol{P}$ に属する測度が特徴付けられるということである. したがって, $\mathcal{G}$ が $\boldsymbol{P}$ に関して十分であれば, $\mathcal{F}$ 全体で論ずる代わりに $\mathcal{G}$ において論ずれば, $\boldsymbol{P}$ に関して完全な情報を得ることができるのである. このことから, 十分性の概念は単に数学的な概念であるにとどまらず自然現象・社会現象の解析に有用な役割を果たすのである. たとえば, この概念が量子力学系の状態の分類に役立つ様を第2巻の第9節で詳しく論じてある.

第4章　種々のエントロピー

この章では，第3章まで論じてきたエントロピー以外の様々なエントロピーについて概説する．

4.1　Rényi エントロピー

任意の確率分布 $p=\{p_k\}\in\Delta_n$ と測度空間 $(\Omega,\mathcal{F},\boldsymbol{p})$ における有限分割 $\tilde{A}\in\mathcal{P}(\mathcal{F})$ に対して，2つの汎函数を

$$S_\alpha(p)=(1-\alpha)^{-1}\log\left(\sum_{k=1}^{n}p_k{}^\alpha\right),$$

$$S_\alpha(\tilde{A})=(1-\alpha)^{-1}\log\left(\sum_{A\in\tilde{A}}p(A)^\alpha\right)$$

と定め，各々 $p,\tilde{A}$ に関する **α 次 Rényi のエントロピー**という．ただし，$\alpha>0$，$\neq 1$ である．

今，$f(\alpha)=\log\sum_{k=1}^{n}p_k{}^\alpha$ とおくと，$\sum_{k=1}^{n}p_k=1$ より $f(1)=0$ であり，さらに，

$$f'(1)=\lim_{\alpha\to1}\frac{f(\alpha)-f(1)}{\alpha-1}$$

$$=\left[\frac{1}{\sum_{k=1}^{n}p_k{}^\alpha}\sum_{k=1}^{n}p_k{}^\alpha\log p_k\right]_{\alpha=1}=\sum_{k=1}^{n}p_k\log p_k$$

であるから，

$$\lim_{\alpha\to1}S_\alpha(p)=S(p)$$

となる．同様にして，

$$\lim_{\alpha\to1}S_\alpha(\tilde{A})=S(\tilde{A}).$$

ここで，2つの確率測度 p,q に対して，**α 次 Rényi の相対エントロピー**を

$$S_\alpha(p|q)=\sup\left\{(\alpha-1)^{-1}\log\left(\sum_{A\in\tilde{A}}\frac{p(A)^\alpha}{q(A)^{\alpha-1}}\right);\quad \tilde{A}\in\mathcal{P}(\mathcal{F})\right\}$$

で定めると，S_α の場合と同様

$$\lim_{\alpha\to 1} S_\alpha(p|q)=S(p|q)$$

が成立する．さらに $p\ll q$ であれば，等式

$$S_\alpha(p|q)=(\alpha-1)^{-1}\log\int\left(\frac{dp}{dq}\right)^\alpha dq$$

も成立する．

ところで，$p\in\Delta_n$ に対して

$$\bar{S}_\alpha(p)=\left(\sum_{k=1}^n p_k{}^\alpha-1\right)\Big/(e^{1-\alpha}-1)$$

とおくと，$S_\alpha(p)$ の定義より

$$\bar{S}_\alpha(p)=\{e^{(1-\alpha)S_\alpha(p)}-1\}/(e^{1-\alpha}-1)$$

となり，さらに，簡単な計算の結果

$$\lim_{\alpha\to 1}\bar{S}_\alpha(p)=S(p)$$

となることがわかる．ここで，$S_\alpha(p)$，$\bar{S}_\alpha(p)$ の主要な性質を列記しておこう．証明は読者自ら試みてほしい．

1°　**正値性**：$S_\alpha(p)\geqq 0$，$\bar{S}_\alpha(p)\geqq 0$.

2°　**拡張可能性**：$p=(p_1,p_2,\cdots,p_n)\in\Delta_n$ に対して $\tilde{p}=(p_1,p_2,\cdots,p_n,0)$ とおくと，$S_\alpha(\tilde{p})=S_\alpha(p)$，$\bar{S}_\alpha(\tilde{p})=\bar{S}_\alpha(p)$.

3°　**対称性**：$p=(p_1,\cdots,p_n)\in\Delta_n$ の p_k を置換したものを $p'=(p_{\pi(1)},\cdots,p_{\pi(n)})$ とすると，$S_\alpha(p')=S_\alpha(p)$，$\bar{S}_\alpha(p')=\bar{S}_\alpha(p)$.

4°　**連続性**：$S_\alpha(p)=S_\alpha(p_1,\cdots,p_n)$，$\bar{S}_\alpha(p)=\bar{S}_\alpha(p_1,\cdots,p_n)$ は $\bigcup_{n\geqq 2}\Delta_n\to\boldsymbol{R}^+$ の連続函数である．

5°　**加法性**：$p\in\Delta_n$，$q\in\Delta_m$ に対して，

$$S_\alpha(p_1q_1,p_1q_2,\cdots,p_nq_m)=S_\alpha(p)+S_\alpha(q).$$

6°　$p\in\Delta_n$ のとき，$\bar{S}_\alpha(p)\leqq\bar{S}_\alpha\left(\frac{1}{n},\frac{1}{n},\cdots,\frac{1}{n}\right)$.

7°　**単調性**：$\bar{S}_\alpha\left(\frac{1}{n},\frac{1}{n},\cdots,\frac{1}{n}\right)$ は n に関して単調増加.

8°　$p_1+p_2>0$，$n\geqq 3$ のとき，

$$\bar{S}_\alpha(p_1, \cdots, p_n) = \bar{S}_\alpha(p_1+p_2, p_3, \cdots, p_n) + (p_1+p_2)^\alpha \bar{S}_\alpha\left(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2}\right).$$

9° $p \in \Delta_n$, $q \in \Delta_m$ に対して

$$\bar{S}_\alpha(p_1q_1, \cdots, p_nq_m) = \bar{S}_\alpha(p) + \bar{S}_\alpha(q) + (e^{1-\alpha}-1)\bar{S}_\alpha(p)\bar{S}_\alpha(q).$$

10° $q_{ij} \geq 0$, $\sum_{j=1}^{n} q_{ij} = 1$ $(i=1, 2, \cdots, m)$, $p \in \Delta_m$ のとき

$$\bar{S}_\alpha(p_1q_{11}, \cdots, p_mq_{mn}) = \bar{S}_\alpha(p) + \sum_{i=1}^{m} p_i{}^\alpha \bar{S}_\alpha(q_{i1}, \cdots, q_{in}).$$

なお，上と同じ結果は，$\tilde{A} \in \mathcal{P}(\mathcal{F})$ に対する Rényi のエントロピーにおいても当然成立する．

4.2 微分エントロピー

n 個の確率変数 $\xi_1, \xi_2, \cdots, \xi_n$ のつくる確率ベクトル $\vec{\xi} = (\xi_1, \xi_2, \cdots, \xi_n)$ の分布が密度函数 $f(t_1, t_2, \cdots, t_n)$ をもつとき

$$S(\vec{\xi}) = -\int f(t_1, t_2, \cdots, t_n) \log f(t_1, t_2, \cdots, t_n) dt_1 \cdots dt_n$$

を確率ベクトル $\vec{\xi}$ の**微分エントロピー**という（ただし，定義されるとき）．この $S(\vec{\xi})$ は $S(f)$ とかくことが多い．

以下，2つの確率ベクトル $\vec{\xi}, \vec{\zeta}$ に対して，$f_\xi(\vec{t})$ を $\vec{\xi}$ の，$f_\zeta(\vec{t})$ を $\vec{\zeta}$ の，$f(\vec{t}, \vec{r})$ を確率ベクトル $(\vec{\xi}, \vec{\zeta})$ の分布の密度函数とする．このとき，Shannon のエントロピーと同様，**条件付微分エントロピー**が次のように定められる．

$$S(\vec{\xi}|\vec{\zeta}) = -\int f(\vec{t}, \vec{r}) \log \frac{f(\vec{t}, \vec{r})}{f_\zeta(\vec{t})} d\vec{t}\, d\vec{r}.$$

これらの微分エントロピーに関して，次の結果が容易に示せる．

1° $-\infty \leqq S(\vec{\xi}) \leqq \infty$；**正値性は成立しない．**

2° $S(\vec{\xi}, \vec{\zeta}) = S(\vec{\xi}) + S(\vec{\zeta}|\vec{\xi}) = S(\vec{\zeta}) + S(\vec{\xi}|\vec{\zeta})$.

3° $S(\vec{\xi}|\vec{\zeta}) \leqq S(\vec{\xi})$.

4° $\vec{\xi}$ と $\vec{\zeta}$ が独立のとき，$S(\vec{\xi}, \vec{\zeta}) = S(\vec{\xi}) + S(\vec{\zeta})$ および $S(\vec{\xi}|\vec{\zeta}) = S(\vec{\xi})$.

5° $\int k(\vec{t}, \vec{r}) d\vec{r} = \int k(\vec{t}, \vec{r}) d\vec{t} = 1$ である $k(\vec{t}, \vec{r}) \geqq 0$ によって

$$f'(\vec{t})=\int k(\vec{t},\vec{r})f(\vec{r})d\vec{r}$$

と密度函数 f を変換すると，$S(f')\geqq S(f)$.

6° $\vec{t}=(t_1,\cdots,t_n)$ をある座標変換によって $\vec{t}'=(t_1',\cdots,t_n')$ に移したとき，$f(\vec{t})$ が $g(\vec{t}')$ に変わるとすると，

$$S(g)=S(f)+\int f(\vec{t})\log|J(\vec{t};\vec{t}')|d\vec{t}.$$

ここで，$J(\vec{t};\vec{t}')$ は変換 $\vec{t}\mapsto\vec{t}'$ に関する Jacobian である．したがって，S は一般に座標軸のとり方に依存している．とくに，$t_j'=\sum_{k=1}^{n}a_{jk}t_k$ のとき，$S(g)=S(f)+\log|\det(a_{ij})|$ となる．よって S は回転に関しては不変性を保っている．

7° **スケール変換**：$\vec{\xi}\mapsto a\vec{\xi}$ $(a>0)$ に関して，

$$S(a\vec{\xi})=S(\vec{\xi})+n\log a$$

となり，やはり不変ではない．

8° 平行移動 $t_j\mapsto t_j'=t_j+a$ に関して $S(\vec{\xi})$ は不変である．

ところで，Boltzmann と Gibbs は N 個の粒子からなる系の相空間 $\boldsymbol{R}^{6N}$ における分布の密度函数を $f(\vec{r})$ $(\vec{r}\in\boldsymbol{R}^{6N})$ とするとき，系のもつエントロピーを

$$S=-\frac{1}{h^{3N}N!}\int f(\vec{r})\log f(\vec{r})d\vec{r}\quad (h \text{ は Planck 定数})$$

と定めた．この S が最初に導入された微分エントロピーである．この S を Ochs[O.1] はさらに一般化した．今 $(\Omega,\mathcal{F})$ を可測空間とし，μ を $\mathcal{F}$ 上の σ-有限測度，$\mathfrak{S}(\mu)$ を μ に関して絶対連続となる有限測度 ν の集合とする．このとき，$\mathfrak{S}(\mu)$ 上の汎函数 S_μ を

$$S_\mu(\nu)=-\int_\Omega \log\left(\frac{d\nu}{d\mu}\right)d\nu$$

と定める．これは，上記の **Boltzmann-Gibbs のエントロピー** S の拡張であること以上に，Shannon のエントロピー，相対エントロピーと次のような関連がある．

（1） $\Omega=\{1,2,\cdots\}$，$\mu(\{k\})=1$，$\nu(\{k\})=p_k\geqq 0$，$\sum p_k=1$ と μ,ν を定めると $S_\mu(\nu)=-\sum p_k\log p_k$.

（2） μ も有限測度とすると $S_\mu(\nu)=-S(\nu|\mu)$.

以上の事実から，$S_\mu(\nu)$ を Ochs [O.1] は **Boltzmann-Gibbs-Shannon のエントロピー**とよんでいる．しかしながら，この $S_\mu(\nu)$ は $S(\vec{\xi})$ と同様，正値性を有しておらず，したがって，$S_\mu(\nu)$ を ν と μ の間の不確定さ，あるいは，情報量を表す量と見なすわけにはいかず，よって，$S_\mu(\nu)$ は力学系の特質を議論するためにはあまり適当なものではないようである．この $S_\mu(\nu)$ に対応する量子力学版が第 2 巻の第 12 章で論ぜられる Segal エントロピーである．

なお，微分エントロピーの典型は **Gauss 分布の密度函数**によるそれである．すなわち，平均 0 のとき

$$f(t_1,\cdots,t_n)=\frac{\sqrt{\det(a_{ij})}}{(\sqrt{2\pi})^n}\exp\left(-\frac{1}{2}\sum_{i,j=1}^{n}a_{ij}t_it_j\right)$$

で

$$S(f)=\log(2\pi e)^{n/2}|\det(a_{ij})|^{-1/2}$$

となる．ここで，行列 (a_{ij}) の逆行列が共分散行列とよばれているものである．

ところで，有限な分散をもつ確率変数 ξ において，その密度函数 f が $L^2(\boldsymbol{R})$ に入るから，$-f\log f\in L^1(\boldsymbol{R})$ となり，$S(\xi)$ が有限確定値をとる．このとき，Gauss 確率変数が次のように特徴付けられる．

〈**定理 4.1**〉 有限な分散 σ^2 をもつ絶対連続な確率変数 ξ のエントロピーに関して，次の関係が成り立つ．

$$S(\xi)\leq\frac{1}{2}\log(2\pi e\sigma^2), \tag{4.1}$$

さらに，ξ が平均 0，分散 σ^2 をもつ Gauss 確率変数のとき，かつそのときに限り等号が成立する．

証明 Lebesgue 測度の平行移動に関する不変性より，ξ の平均は 0 としてよい．今，同じ σ^2 を分散にもつ平均 0 の Gauss 確率変数を ζ，その密度函数を g とすると，

$$g(t)=(2\pi\sigma^2)^{-1/2}\exp(-t^2/2\sigma^2)$$

であり，次の積分の計算を行うと

$$-\int f(t)\log g(t)dt=\int f(t)\left\{\frac{1}{2}\log(2\pi\sigma^2)+\frac{t^2}{2\sigma^2}\right\}dt$$
$$=\frac{1}{2}\log(2\pi\sigma^2)+\frac{1}{2\sigma^2}\int t^2 f(t)dt$$
$$=\frac{1}{2}\log(2\pi\sigma^2)+\frac{1}{2}=\frac{1}{2}\log(2\pi e\sigma^2).$$

したがって，〈補題 3.2〉を用いて(4.1)を得る．これが等式となることと $f(t)=g(t)$ a.e. とは同値．▌

上の結果を用いると，有限な分散をもつ絶対連続な確率変数 ξ と Gauss 確率変数 ζ の Kullback-Leibler の情報量 $I(\xi|\zeta)$ は，ξ が Gauss 型のときに最小となることがわかる．実際，ξ と ζ の密度函数を f, g とすると，

$$I(\xi|\zeta)=-S(\xi)-\int_R f(t)\log g(t)dt$$

となる．ここで，右辺の第二項は，ξ の分布によらず一定であるが，第一項は，ξ が Gauss 型のときに最小となるからである．

4.3 ε-エントロピー

距離空間 $X=(X,d)$ において，部分集合 $Y\subset X$ の直径を $D(Y)=\sup\{d(x,y); x,y\in Y\}$ とおく．X の部分集合の族 $\mathcal{U}=\{U,V,\cdots\}$ が $Y\subset X$ の **ε-被覆**であるとは

$$Y\subset\cup\{U; U\in\mathcal{U}\},\qquad \sup\{D(U); U\in\mathcal{U}\}\leqq 2\varepsilon$$

のときをいい，部分集合 $W\subset X$ が Y の **ε-網**であるとは，任意の $y\in Y$ に対して $w\in W$ が存在して $d(w,y)<\varepsilon$ を満たすときをいう．また，集合 $Y\subset X$ が **ε-分離**であるとは，任意の対 $a,b\in Y$ $(a\neq b)$ がつねに $d(a,b)>\varepsilon$ であるときをいう．このとき次の定理が成り立つ．

〈定理 4.2〉 距離空間 X の部分集合 Y に対して，次の4つの条件はたがいに同値である．

（1） 任意の $\varepsilon>0$ に対して，Y の有限な ε-被覆が存在する．

（2） 任意の $\varepsilon>0$ に対して，Y の有限な ε-網が存在する．

（3） 任意の$\varepsilon>0$に対して，Yに含まれるε-分離集合はつねに有限である.

（4） Yは全有界（＝プレコンパクト）である.

註 これらの証明は距離空間における基本的手法で行うことができる．これらに関連てし，Yが完備距離空間ならば，部分集合$Y\subset X$に対して次の関係が成り立つ：Y上の実数値連続函数は最大値，最小値をもつ$\Longleftrightarrow Y$はコンパクト$\Longleftrightarrow Y$は閉かつ全有界$\Rightarrow Y$は全有界$\Rightarrow Y$ は可分（すなわち，可算稠密集合が存在する).

したがって，空間 X 自身が全有界またはコンパクトな場合は〈定理 4.2〉が適用される．典型的な例はメッセージ空間 $X=A^Z$ である（次章参照).

これらの概念の数量表示が，次のように定義される.

[定義 4.1] 部分集合 $Y\subset X$ に対して，

（1） $N(Y,\varepsilon)$ は Y の ε-被覆を構成する X の部分集合の個数の最小値;

（2） $N_X(Y,\varepsilon)$ は Y の ε-網を構成する X の元の個数の最小値;

（3） $M(Y,\varepsilon)$ は Y の ε-分離集合を構成する Y の元の個数の最大値;

（4） $S(Y,\varepsilon)\equiv\log N(Y,\varepsilon)$: Y の **ε-エントロピー**;

（5） $S_X(Y,\varepsilon)\equiv\log N_X(Y,\varepsilon)$: Y の **X における ε-エントロピー**;

（6） $C(Y,\varepsilon)\equiv\log M(Y,\varepsilon)$: Y の **ε-キャパシティ**.

ここで，ε-網を用いて定義する $N_X(Y,\varepsilon)$ と $S_X(Y,\varepsilon)$ は Yを含む距離空間 X の選び方に従属して定まる．他は Y のみに依存する.

これら5種類の集合函数の基本的性質について概説する.

〈定理 4.3〉 全有界部分集合 $Y\subset X$ に対して，つねに次の不等式が成立する.

（1） $M(Y,2\varepsilon)\leqq N(Y,\varepsilon)\leqq N_X(Y,\varepsilon)\leqq N_W(Y,\varepsilon)\leqq M(Y,\varepsilon)$.

（2） $C(Y,2\varepsilon)\leqq S(Y,\varepsilon)\leqq S_X(Y,\varepsilon)\leqq S_W(Y,\varepsilon)\leqq C(Y,\varepsilon)$.

ここで，W は任意の $Y\subset W\subset X$ となる集合で，$\varepsilon>0$ も任意である.

証明 （1）の最も右側と最も左側の2つの不等式を示す．他は容易である．$\{a_1,\cdots,a_n\}\subset Y$ $(n\equiv M(Y,\varepsilon))$ が Y の ε-分離集合ならば，ε-網でもある．し

たがって，W に含まれる ε-網の最小値が $N_W(Y,\varepsilon)$ であるから目的の最右辺の不等式がいえる．最左辺の不等式の証明は，もし不等号の向きが逆ならば

$$[2\varepsilon\text{-分離集合を構成する } Y \text{ の元の個数}] \gneqq [\varepsilon\text{-被覆の集合 } \{V_k\} \text{ の個数}]$$

となる．よって，ある V_k と $a, b \in V_k$ が存在して $d(a,b) > 2\varepsilon$ となる．ところが $a, b \in V_k$ であるから $d(a,b) \leqq 2\varepsilon$. これは矛盾である．これで（1）が示された．（2）は（1）の帰結である．▌

一般に，X の各部分集合 Y および Y_k に対して正値集合函数 $F(Y)$ が定義されていて，つねに

$$Y \subset \bigcup_{k=1}^{n} Y_k \Longrightarrow F(Y) \leqq \sum_{k=1}^{n} F(Y_k)$$

であるとき，$F(\cdot)$ を集合函数として**半加法的**とよぶ．これについて，次の定理が容易に示せる：

〈**定理 4.4**〉 6種類の函数 $N(Y,\varepsilon)$, $N_X(Y,\varepsilon)$, $M(Y,\varepsilon)$, $S(Y,\varepsilon)$, $S_X(Y,\varepsilon)$, $C(Y,\varepsilon)$ について，これらは

（1） 集合函数として半加法的である；

（2） $Y \subset X$ が全有界部分集合とするとき，上記の6種類の函数は ε を変数と考えるとき，減少函数であり，さらに $N(Y,\varepsilon)$, $M(Y,\varepsilon)$, $S(Y,\varepsilon)$, $C(Y,\varepsilon)$ は右連続函数である．

〈**定理 4.5**〉 $N(Y,\varepsilon) = \min\limits_{Y \subset W} N_W(Y,\varepsilon)$, $S(Y,\varepsilon) = \min\limits_{Y \subset W} S_W(Y,\varepsilon)$. ここで，$W$ は Y を部分空間として含む距離空間である．

この "min" による等式の帰結として $S(Y,\varepsilon)$ を**最小 ε-エントロピー**とよぶ．

上述の各数量のうち，重要な $N_X(Y,\varepsilon)$, $S_X(Y,\varepsilon)$ の情報理論上の意味を考察しよう．X をアルファベット・メッセージ空間 A^Z とすると，これはコンパクト距離空間となる．この A^Z 上で種々な通報の解析がなされる．（第5章参照）．$Y \subset A^Z$ をある通報の集団とする．このとき $y \in Y$ が ε (>0) の精度で**実現可能**であるとは，y との距離が ε より小さい既知の $y_0 \in A^Z$ が存在すること

である（これは通報を高精度で認識する方法のひとつの数学的な表現形式と考えられる）．この方法によって，未知の通報 $y \in Y$ を解読するためには，$N_X(Y, \varepsilon)$ 個（$X = A^Z$）の相異なる既知の通報を用いれば ε の精度で目的が達成される．ε-エントロピーについての解釈は次のとおりである：情報理論における単位のとり方は，基本的に "bit" を採用している（第1章の 1.3 節参照）．このとき，上記のメッセージ空間 A^Z におけるアルファベット A は通常 $A = \{0, 1\}$ と二点集合にとれる．そこで観測の対象となる通報の集団 $Y \subset A^Z$ に対して，十分小さな $\varepsilon > 0$ を与えたとき，各通報 $y \in Y$ に長さ $h = [S_X(Y, \varepsilon)]$（ただし，$[\cdot]$ は Gauss 記号）の二価記号（0または1）からなる既知の通報の列が対応して，y が ε の精度で観測者にもたらされる．

次に具体的な例を2つ挙げる．

1° X が $X \subset \boldsymbol{R}^n$ でコンパクトなとき，X の内点が空集合でなければ

$$N(X, \varepsilon) \sim \varepsilon^{-n} \quad かつ \quad S(X, \varepsilon) \sim -n \log \varepsilon + O(1).$$

2° X が n 次元平行体 (parallelopiped) のとき，$|X|$ を X の体積とすると，

$$N(X, \varepsilon) \sim c_n |X| \varepsilon^{-n} \quad かつ \quad M(X, \varepsilon) \sim d_n |X| \varepsilon^{-n}.$$

ここで，c_n, d_n は n に依存するある定数である．

Hilbert 第 13 問題：ε-エントロピー理論は種々の函数空間上において詳細に展開されてきたが，そのひとつの成果として Arnold と Kolmogorov [K. 11] が 1957 年にこの大問題を画期的な形で解決した．一般に多変数の函数がより少ない変数をもついくつかの函数の結合で表されるとき，これを**函数の函数による重ね合わせ**，または**合成**という．たとえば，実数値3変数函数 φ に対して2変数函数 f, g, h, k が存在して

$$\varphi(x, y, z) = f(g(x, h(y, z)), \ k(x, y))$$

の場合，φ は f, g, h, k の合成という．より簡単な例としては

$$x_1 x_2 \cdots x_n = x_1(x_2(\cdots(x_{n-1} x_n)\cdots)).$$

Hilbert 第 13 問題は「3変数の連続函数は必ずしも2変数の連続函数の合成によって表すことはできない」というものであり，これを Arnold と Kolmogorov

が否定的に解いたのである．それは，より一般的に，次の定理の形で述べることができる．

〈**定理 4.6**〉　**Kolmogorov の定理**：閉区間 $[0,1]$ に値をとる $n(2n+1)$ 個の単調増加函数 $f_{ij}\in C[0,1]$ $(i=1,\cdots,n;\ j=1,\cdots,2n+1)$ が存在して，すべての n 変数函数 $\varphi\in C([0,1]^n)$ は次の合成式によって表示される：

$$\varphi(x_1,\cdots,x_n)=\sum_{j=1}^{2n+1}\psi\left(\sum_{i=1}^{n}f_{ij}(x_i)\right),$$

ここで $\psi\in C[0,n]$ は φ のみに従属して定まり，f_{ij} はたがいに独立である．

4.4　位相エントロピー

位相エントロピーは，Adler [A.5] などによって 1963 年に導入された新しい概念である．これは，対象とする空間を被覆することによって定義された ε-エントロピーに発想の源があり，しかもその構成は，空間の可測分割に保側変換を作用させる Shannon–Kolmogorov–Sinai 流のエントロピーのアナロジーである．また，このエントロピーは位相空間上の写像に新たな数量を対応させ，写像の位相的性質を論ずる手がかりを与えている．

以下考察する空間はコンパクトであると仮定する．

上述のように，情報理論的な構成法を用いるので，第2章の 2.3 節における K·S エントロピー，および第5章の 5.2 節で考察する情報源エントロピーと同様の方法で，位相エントロピーは扱われる．したがって，以下の解説にもこれらの章の議論が深い係わりをもっている．

この節では X をコンパクト Hausdorff 空間，$\mathcal{C},\mathcal{D},\cdots$ などを X の開被覆，$\mathcal{C}$ の有限部分被覆の全体を $\boldsymbol{F}(\mathcal{C})$，各 $\mathcal{C}'\in\boldsymbol{F}(\mathcal{C})$ の要素の個数を $|\mathcal{C}'|$ とすると，

$$N(\mathcal{C})\equiv\min\{|\mathcal{C}'|;\ \mathcal{C}'\in\boldsymbol{F}(\mathcal{C})\}$$

はつねに存在し，この対数をとった

$$S(\mathcal{C})\equiv\log N(\mathcal{C})$$

を $\mathcal{C}$ の**位相エントロピー**という．これは X の開被覆 $\mathcal{C}$ に対して一意に定まる量であり，確率空間の有限可測分割によるエントロピーと並行した議論がなされる．2つの開被覆 $\mathcal{C},\mathcal{D}$ の間に次の記号 $\prec$, $\vee$ を定義する；

$\mathcal{C}\prec\mathcal{D}\overset{\text{def}}{\Longleftrightarrow}$ 各 $D\in\mathcal{D}$ はいずれかの $C\in\mathcal{C}$ の部分集合．

$\mathcal{C}\vee\mathcal{D}\equiv\{C\cap D;\ C\in\mathcal{C},\ D\in\mathcal{D}\}$.

明らかに $\prec$ によって，すべての開被覆の族 $\mathcal{F}(X)$ は（共終な）順序集合となり，$\mathcal{F}(X)$ 上の非負な実数値函数 $N(\cdot)$, $S(\cdot)$ は次の性質を満足する．

1° $\mathcal{C}\prec\mathcal{D}\Longrightarrow N(\mathcal{C})\leqq N(\mathcal{D})$, $S(\mathcal{C})\leqq S(\mathcal{D})$,

2° $N(\mathcal{C}\vee\mathcal{D})\leqq N(\mathcal{C})N(\mathcal{D})$, $S(\mathcal{C}\vee\mathcal{D})\leqq S(\mathcal{C})+S(\mathcal{D})$.

T が $X\to X$ なる連続写像であるとき，開被覆 $\mathcal{C}$ に対して $T^{-1}\mathcal{C}=\{T^{-1}C;\ C\in\mathcal{C}\}$ も開被覆であり，

3° $N(T^{-1}\mathcal{C})\leqq N(\mathcal{C})$，ここで T が自己同型ならば等号が成立．

以下，定常情報源の (Shannon) エントロピー $S(\mu)$ (第5章5.2節参照) や，確率測度空間上の保測変換の (Kolmogorov–Sinai) エントロピー（第2章，2.3節参照）などと同様な構成により，連続写像のエントロピーが導入される．T を連続写像，$\mathcal{C}$ を開被覆とし，

$$\bigvee_{k=0}^{n-1}T^{-k}\mathcal{C}\equiv\mathcal{C}\vee T^{-1}\mathcal{C}\vee\cdots\vee T^{-n+1}\mathcal{C},$$

$$S_n(T,\mathcal{C})\equiv S\left(\bigvee_{k=0}^{n-1}T^{-k}\mathcal{C}\right)$$

とおくと，

4° X 上の連続写像 T と開被覆 $\mathcal{C}$ に対して

$$S(T,\mathcal{C})\equiv\lim_{n\to\infty}\frac{1}{n}S_n(T,\mathcal{C})$$

がつねに存在する．

これは第2章の2.3節の (2.5) のそれと同じであるが，次のように証明することもできる．

略証 次の不等式を使い，あとは簡単な計算により示せる（第5章でも同様な議論がある）．

$$
\begin{aligned}
S_{m+n}(T,\mathcal{C}) &= S\left(\bigvee_{k=0}^{m+n-1} T^{-k}\mathcal{C}\right)\\
&= S\left(\left(\bigvee_{k=0}^{m-1} T^{-k}\mathcal{C}\right)\vee T^{-m}\left(\bigvee_{k=0}^{n-1} T^{-k}\mathcal{C}\right)\right)\\
&\leqq S\left(\bigvee_{k=0}^{m-1} T^{-k}\mathcal{C}\right)+S\left(T^{-m}\bigvee_{k=0}^{n-1} T^{-k}\mathcal{C}\right)\\
&\leqq S\left(\bigvee_{k=0}^{m-1} T^{-k}\mathcal{C}\right)+S\left(\bigvee_{k=0}^{n-1} T^{-k}\mathcal{C}\right)\\
&= S_m(T,\mathcal{C})+S_n(T,\mathcal{C}). \quad \blacksquare
\end{aligned}
$$

[定義 4.2]　この定理によって得られた $S(T,\mathcal{C})$ を $\mathcal{C}$ による T の位相エントロピーといい，さらに

$$S_T \equiv \sup\{S(T,\mathcal{C}) \mid \mathcal{C}\in\mathcal{F}(X)\}$$

を $\boldsymbol{T}$ **の位相エントロピー**という．

この位相エントロピー S_T と力学系 $(X,\mathcal{F},\mu,T)$ に関する Kolmogorov-Sinai エントロピー $S_\mu(T)$（ここでは μ を陽に表示した．第2章では $S(T)$ と記してある）との関係は，次の命題によって表される．

5°　X はコンパクト距離空間，T は $X\to X$ である連続写像，$\boldsymbol{P}_T(X)$ を T-不変な X 上の Borel 確率測度の全体とすると

$$S_T=\sup\{S_\mu(T);\ \mu\in\boldsymbol{P}_T(X)\}.$$

6°　X はコンパクト位相群で第一可算公理を満たし（したがって，X は距離群），T は X 上の自己同型とする．また，μ を X 上の Haar 測度で $\mu(X)=1$ とする．このとき μ は T-不変で，$S_T=S_\mu(T)$ を満足する．

これらのことは，数学の様々な分野にも適用され，たとえば，種々の多様体上で位相写像に対するエントロピーが計算され，それを使って力学系の理論が展開されている．

4.5 fuzzy エントロピー

Fuzzy（ファジィー）集合論は Zadeh [Z.1] が 1965 年に導入し，その後多くの研究者によって論じられているが，そのうちとくに De Luca と Termini [D.1, 2] や Knopemacher [K.9] などによって厳密な数学理論として定式化され，興味ある分野になりつつある．元来，この概念は種々な事象について実験したり，測定したり，また図形とか物体の形状を分析，観察したりする場合にどうしても付きまとうあいまいさの度合いを数量表示して，ぼやけ現象を理論的に把握することを目的としたものである．

X を与えられたひとつの集合とし，X の部分集合の全体 2^X と X から $\{0, 1\}$ への函数の全体 $\{0,1\}^X$ とは，$A \to 1_A$（ここで，部分集合 $A \subset X$ に対して，1_A は $X \to \{0,1\}$ の特性函数である）によって，1 対 1 に対応している．ここで，X のひとつの部分集合を考えるということは，X の上に，ある条件（とか性質）が設定されて，その条件（とか性質）を満たす X の要素の全体を考えることで，言い換えれば，“X の要素 x がその部分集合に属する” $\Longleftrightarrow$ “その条件を x が満たす” ということである．部分集合を A，条件を P とすると，“$x \in A$” $\Longleftrightarrow$ “x は P を満足する”．したがって，この A は $A = \{x \in X; x$ は P を満足する$\}$ と表せる．これを特性函数 1_A で表示すれば，“$x \in A$” $\Longleftrightarrow$ “$1_A(x) = 1$”．これらの命題の対偶をとれば，“$x \notin A$” $\Longleftrightarrow$ “x は P を満足しない” $\Longleftrightarrow$ “$1_A(x) = 0$”．つまり，X の要素はある集合に属すか属さないか，条件を満足するかしないか，真か偽かのいずれかによって完全に決定され，あいまいな要素は何も含まれない．一方，この節のはじめに述べたように，図形や物体の形状をできるだけ正確に認識することを考えるとき，“できるだけ正確に” ということ自体に，すでにあいまいさやぼやけが存在してしまっている．100% は確実でないことが多いのである．

今，X をひとつの集合とする．これは特定の平面や空間と思ってもよい．その要素のある集団を認識しようとする．その際，その目標の集団の実体を完全には把握しえない場合を考える（実際上は多くの場合がこうである）．これは，

境界が明確でないものや，幾何的に錯綜した図形によって囲まれている集合，あるいは，不確定な条件によって定義付けられた要素の集まりなどである．その認識しようとする要素の集団を A とする．これは，いままで論じた X の通常の部分集合とは異なり，特性函数 $1_A: X\to\{0,1\}$ には対応できない．そこで，要素 $x\in X$ が A に属するかどうかを表す正確さの程度を数で表示し，それを $f_A(x)$ とおく．一般に，$0\leqslant f_A(x)\leqslant 1$ とし，これが1に近ければ近いほど，$x\in X$ が A に属する度合いが大で，また $f_A(x)$ が0に近いほど，属する度合いが小（つまり，属さない度合いが大）とする．これらの場合は，あいまいさが少ないともいえる．また $f_A(x)$ が 1/2 に近いほど，あいまいさが大という具合いに，函数 $f_A(x)$ を特徴付ける．このようにして，対 (A, f_A) を考え，A を X の **fuzzy 部分集合**，または，単に **fuzzy 集合**といい，f_A を A の**表示函数**（menbership function）という．

以上の説明でわかるように，A は通常の部分集合としての実体をもたない場合が考えられ，これは X 上の漠然としたある概念を表している場合である．

$\boldsymbol{F}(X)$ を $f: X\to[0,1]$ である函数 f 全体の集合とする．任意の $f\in\boldsymbol{F}(X)$ は，X のある fuzzy 集合 A の表示函数 $f=f_A$ と考えられ，fuzzy 集合を考えることは，すなわちいずれかの $f\in\boldsymbol{F}(X)$ を考えることになるのである．$f_A(x)$ の値が1に近いほど，x は fuzzy 集合 A に関係が深いということである．

ここで，fuzzy 集合演算を定義する．

［定義 4.3］ $\mathcal{F}(X)$ を X の fuzzy 集合の全体の族とする．任意の A, A_i $(i\in J), B, C\in\mathcal{F}(X)$ に対して

（1）　$A=B \overset{\text{def}}{\Longleftrightarrow} f_A(x)=f_B(x)$,

（2）　$A=\emptyset \overset{\text{def}}{\Longleftrightarrow} f_A(x)=0$,

（3）　$A=X \overset{\text{def}}{\Longleftrightarrow} f_A(x)=1$,

（4）　$A=B^c \overset{\text{def}}{\Longleftrightarrow} f_A(x)=1-f_B(x)$,

（5） $A \subset B \overset{\text{def}}{\Longleftrightarrow} f_A(x) \leq f_B(x)$,

（6） $A = \cup A_i \overset{\text{def}}{\Longleftrightarrow} f_A(x) = \sup f_{A_i}(x)$,

（7） $A = \cap A_i \overset{\text{def}}{\Longleftrightarrow} f_A(x) = \inf f_{A_i}(x)$,

（6′） $A = B \cup C \overset{\text{def}}{\Longleftrightarrow} f_A(x) = \max(f_A(x), f_B(x))$,

（7′） $A = B \cap C \overset{\text{def}}{\Longleftrightarrow} f_A(x) = \min(f_A(x), f_B(x))$.

ここで，これらのうち函数 f_A, f_B などに関する等式や不等式は，いうまでもなくすべての $x \in X$ についての意味である．（1）から容易に

（1′） $A \neq B \Longleftrightarrow$ いずれかの $x \in X$ で $f_A(x) \neq f_B(x)$ がいえる．

また，（2）の A を **fuzzy 空集合**，（4）の A を B の **fuzzy 補集合**という．（5）のとき，A は B の **fuzzy 部分集合**といい，A は B に含まれ，B は A を含むという．

次の関係式は通常の集合の場合と同様な式であり，これらは定義から容易に示せる：

$$A \cup B = B \cup A, \qquad A \cap B = B \cap A;$$
$$A \cup A = A, \qquad A \cup \emptyset = A, \qquad A \cup X = X;$$
$$A \cap A = A, \qquad A \cap \emptyset = \emptyset, \qquad A \cap X = A;$$
$$A \subset B \ \text{かつ}\ B \subset A \Longleftrightarrow A = B;$$
$$A \subset B \ \text{かつ}\ B \subset C \Longrightarrow A \subset C;$$

ただし，$A \cap A^c = \emptyset$，$A \cup A^c = X$ は一般には成立しない．さらに，次のような fuzzy 集合演算が成立する：

1° $A \cup B$ は A と B を含む最小の fuzzy 集合．

実際，$A, B \subset A \cup B$ は明らか．$A \subset C, B \subset C$ とすると $f_A(x) \leq f_C(x)$，$f_B(x) \leq f_C(x)$．ゆえに，

$$f_{A \cup B}(x) = \max(f_A(x), f_B(x)) \leq f_C(x).$$

すなわち，$A \cup B \subset C$，ゆえに $A \cup B$ の最小性がでる．

2° $A \cap B$ は A, B の双方に含まれる最大の fuzzy 集合．この証明も容易．

さらに,

3°　$A\cup(B\cup C)=(A\cup B)\cup C$,

4°　$A\cap(B\cap C)=(A\cap B)\cap C$.

これらは，表示函数の値の演算“max”と“min”を行うことによっていえる(容易である).

5°　$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$,

6°　$A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$.

5°の証明　次の等式を示せば十分である．$\forall x\in X$ に対して,

$$\min\{f_A(x),\max(f_B(x),f_C(x))\}$$
$$=\max\{\min(f_A(x),f_B(x)),\ \min(f_A(x),f_C(x))\}.$$

ここで，$x\in X$ を任意に選んで固定する．$a=f_A(x)$, $b=f_B(x)$, $c=f_C(x)$ とおくと，上の等式は,

$$a\wedge(b\vee c)=(a\wedge b)\vee(a\wedge c)$$

となる，ただし $a\vee b=\max(a,b)$, $a\wedge b=\min(a,b)$. この等式は，単純な実数の大小関係からすぐいえる．▌

7°　$A^{cc}(\equiv(A^c)^c)=A$,

8°　$A\subset B\Longleftrightarrow A^c\supset B^c$,

9°　$(\cup A_\iota)^c=\cap A_\iota{}^c$,

10°　$(\cap A_\iota)^c=\cup A_\iota{}^c$.

証明　7°：$f_{A^{cc}}(x)=1-(1-f_A(x))=f_A(x)$.

8°は自明.

9°：$f_{(\cup A_\iota)^c}(x)=1-f_{\cup A_\iota}(x)=1-\sup f_{A_\iota}(x)$

$$=1+\inf\{-f_{A_\iota}(x)\}=\inf\{1-f_{A_\iota}(x)\}$$
$$=\inf f_{A_\iota{}^c}(x)=f_{\cap A_\iota{}^c}(x).\ ▌$$

X の fuzzy 集合の全体の族 $\mathcal{F}(X)=\{A,B,\cdots\}$ に $\cup,\cap,{}^c$ が定義されたが，これによって $\mathcal{F}(X)$ は完備な束となる．したがって，表示函数全体の族 $\boldsymbol{F}(X)$ も $f_A\vee f_B=f_{A\cup B}$, $f_A\wedge f_B=f_{A\cap B}$, $f_A{}^c=f_{A^c}$ によって完備な束となり，こうして $f\vee g$, $f\wedge g$, f^c が定義される．また任意の $A,B\in\mathcal{F}(X)$ と任意の α,β

$(\alpha+\beta=1,\ \alpha,\beta\geq 0)$ に対して $\alpha A+\beta B$ を $f_{\alpha A+\beta B}=\alpha f_A+\beta f_B$ によって定義すると，$\alpha A+\beta B$ は fuzzy 集合であり，$\mathcal{F}(X)$ は凸集合となる．さらに任意の $A,B\in\mathcal{F}(X)$ に対して積 $A\circ B$ を $f_{A\circ B}=f_A f_B$ で定義すると，$\mathcal{F}(X)$ は可換な半群となる．これらの数学構造のもとで，次のことが成立する．

〈定理 4.7〉 次の条件はたがいに同値である．

（1） $A\in\mathcal{F}(X)$ が X の通常の部分集合（$\subset X$）である．

（2） $A\in\mathcal{F}(X)$ は凸集合として $\mathcal{F}(X)$ の端点である（すなわち，A は $\mathcal{F}(X)$ の他の元の凸結合では表わされない元）．

（3） $A\in\mathcal{F}(X)$ は半群として $\mathcal{F}(X)$ のべき等元である (i.e., $A^2=A$)．

今，表示函数全体の族 $\boldsymbol{F}(X)$ を定義域とする汎函数 $D(\cdot)$ を考え，これに対して，次の条件〈D.1〉～〈D.6〉を満たすとき，$D(f)$ を $f\in\boldsymbol{F}(X)$ の **fuzzy エントロピー**と名付ける：

〈D.1〉 $f(x)=0$ または 1（すなわち f は $\{0,1\}$ の 2-値函数）$\Longleftrightarrow D(f)=0$．

〈D.2〉 $f(x)\equiv\dfrac{1}{2}\Longleftrightarrow D(f)=\max\{D(g);\ g\in\boldsymbol{F}(X)\}$．

［定義 4.4］ $$f\prec g \overset{\text{def}}{\Longleftrightarrow} \begin{cases} f(x)\geqq g(x) & (g(x)\geqq 1/2 \text{ のとき}), \\ f(x)\leqq g(x) & (g(x)\leqq 1/2 \text{ のとき}). \end{cases}$$

このとき f は g よりシャープまたはぼやけが小であるという．

この記号のもとで，

〈D.3〉 $f\prec g\Longrightarrow D(f)\leqq D(g)$．

〈D.4〉 $D(f)=D(f^c)$，$\forall f\in\boldsymbol{F}(X)$ $(f^c=1-f)$．

〈D.5〉 $D(\cdot)$ は束 $\boldsymbol{F}(X)$ 上の非負変分性をもつ：

$$D(f\vee g)+D(f\wedge g)=D(f)+D(g).$$

〈D.6〉 $D(\cdot)$ は $\boldsymbol{F}(X)$ の距離 ρ に関して連続，ここで $\rho(f,g)\equiv\sup\{|f(x)-g(x)|;\ x\in X\}$．

各 t $(0\leqq t\leqq 1/2)$ に対して $\varphi_t(x)\equiv t$, $\forall x\in X$ とおくと $\varphi_t\in \boldsymbol{F}(X)$ であり，これに対してさらに $\varDelta(t)\equiv D(\varphi_t)$ とおき，次の条件 ⟨$\varDelta$.1⟩ を考える：

⟨**$\varDelta$.1**⟩ 閉区間 $[0, 1/2]$ 上で $\varDelta(t)\uparrow\uparrow$（狭義単調増加).

函数 $\varDelta(\cdot)$ については $1/2<t\leqq 1$ のとき $\varDelta(t)$ を $\varDelta(t)\equiv\varDelta(1-t)$ と定義すると，$\varDelta(t)$ の定義域は $[0, 1]$ となり，

⟨**$\varDelta$.2**⟩ $\varDelta(0)=\varDelta(1)=0$, $\varDelta(t)=\varDelta(1-t)$ $(t\in[0,1])$ を満たす.

この fuzzy エントロピー（以下単にエントロピー）$D(f)$ についての各条件を分析してみる．まず ⟨D.1⟩ は，A のエントロピーが 0 ということは fuzzy 集合 A のあいまい度が全くないということで，Shannon エントロピーの場合：$S(p)=0\Longleftrightarrow p_k=1$, $p_j=0$ $(j\neq k)$ に対応し，A がちょうど通常の集合であるという主張である．⟨D.2⟩ は，A のエントロピーが最大なのは表示函数が恒等的に 1/2 のときであり，Shannon エントロピーの条件 ⟨S.3⟩ に対応．⟨D.3⟩ は表示函数が 1/2 に近いほどエントロピーが大ということ．⟨D.4⟩ はエントロピーの対称性を表し，⟨D.5⟩ は後で述べるように Shannon エントロピーと同様，エントロピー函数の次元函数としての特性を表している．⟨D.6⟩ は Shannon エントロピーの連続性の条件 ⟨S.2⟩ に対応している．

ここで $X=\{x_1, x_2, \cdots, x_n\}$（有限集合）とし，各 $f\in\boldsymbol{F}(X)$ に対して Shannon エントロピーの形式で

$$S(f)\equiv-\sum_{i=1}^{n} f(x_i)\log f(x_i)$$

を考える．$S(\cdot)$ は条件 ⟨D.1⟩, ⟨D.5⟩, ⟨D.6⟩ を満足する（証明は容易)．この $S(f)$ を用いて，

$$D(f)=S(f)+S(1-f)$$

とおくと，これはすべての条件 ⟨D.1⟩～⟨D.6⟩ を満足する．

これらは有限集合の場合であったが，ここで無限，連続集合ともいうべき場合を含む X について有用な定理を述べよう．

$(X, \mathcal{F}_X, \mu)$ を確率測度空間，$\boldsymbol{F}(X, \mathcal{F}_X)\equiv\{f\in\boldsymbol{F}(X)\,;\,f$ は $\mathcal{F}_X$-可測$\}$ とするとき [K.9],

〈**定理 4.8**〉 閉区間 $[0,1]$ 上の実数値函数 $\Delta(\cdot)$ が条件〈$\Delta.1$〉と〈$\Delta.2$〉を満たすならば，$\boldsymbol{F}(X,\mathcal{F}_X)$ 上の汎函数

$$D(f)=\int_X \Delta(f(x))d\mu(x)$$

は条件〈D.1〉～〈D.6〉を満足し，$\Delta(t)=D(\varphi_t)$ となる．ただし，〈D.1〉～〈D.4〉における $\boldsymbol{F}(X)$ に属する函数についての等号，不等号はすべて μ-a.e. の意味である．

証明 〈D.1〉: $D(f)=0 \Longleftrightarrow \Delta(f(x))=0$ μ-a.e. $\Longleftrightarrow f(x)=0$ または 1 μ-a.e..

〈D.3〉: $f \prec g$ のとき，$X_g \equiv \{x\in X;\ g(x)\leqq 1/2\}$ とおくと $f(x)\leqq g(x)\leqq 1/2$ $(\forall x\in X_g)$，$1/2\leqq g(x)\leqq f(x)$ $(\forall x\in X_g{}^c)$．ゆえに，

$$D(f)=\left(\int_{X_g}+\int_{X_g{}^c}\right)\Delta(f(x))d\mu(x)$$
$$\leqq\left(\int_{X_g}+\int_{X_g{}^c}\right)\Delta(g(x))d\mu(x)=D(g).$$

〈D.2〉: 〈D.3〉から $D(f)\leqq D(1/2)$．ゆえに $D(f)=D(1/2)\Longrightarrow f\equiv 1/2$ をいえばよい．一般に，$\Delta(1/2)-\Delta(f(x))\geqq 0$ であるが，$\Delta(\cdot)$ が $[0,1/2]$ 上で狭義単調増加であるから，

$$0=D(1/2)-D(f)=\int_X(\Delta(1/2)-\Delta(f(x)))d\mu(x)$$

$$\Rightarrow \Delta(1/2)-\Delta(f(x))=0\ \mu\text{-a.e.} \Rightarrow f(x)=\frac{1}{2}\ \mu\text{-a.e..}$$

〈D.4〉: $\Delta(f^c(x))=\Delta(1-f(x))=\Delta(f(x))$.

〈D.5〉: $h\equiv f-g$ は X 上で実数値であり

$$\left.\begin{matrix} f\vee g=f \\ f\wedge g=g \end{matrix}\right\}(h^{-1}[0,\infty)\ 上),\qquad \left.\begin{matrix} f\vee g=g \\ f\wedge g=f \end{matrix}\right\}(h^{-1}(-\infty,0)\ 上).$$

これから

$$D(f\vee g)+D(f\wedge g)$$
$$=\int_X \Delta(f\vee g(x))d\mu(x)+\int_X \Delta(f\wedge g(x))d\mu(x)$$

$$=\left(\int_{h^{-1}[0,\infty)}+\int_{h^{-1}(-\infty,0)}\right)\Delta(f\vee g(x))\,d\mu(x)$$
$$+\left(\int_{h^{-1}[0,\infty)}+\int_{h^{-1}(-\infty,0)}\right)\Delta(f\wedge g(x))\,d\mu(x)$$
$$=\int_{h^{-1}[0,\infty)}\Delta(f(x))\,d\mu(x)+\int_{h^{-1}(-\infty,0)}\Delta(g(x))\,d\mu(x)$$
$$+\int_{h^{-1}[0,\infty)}\Delta(g(x))\,d\mu(x)+\int_{h^{-1}(-\infty,0)}\Delta(f(x))\,d\mu(x)$$
$$=\int_X\Delta(f(x))\,d\mu(x)+\int_X\Delta(g(x))\,d\mu(x)$$
$$=D(f)+D(g).$$

〈D.6〉: $\rho(f_n,f)\to 0 \Rightarrow f_n(x)\to f(x)$ （一様収束）

$$\Rightarrow \Delta(f_n(x))\to\Delta(f(x))\ \text{（一様収束）}$$
$$\Rightarrow D(f_n)=\int\Delta(f_n(x))\,d\mu(x)$$
$$\to\int\Delta(f(x))\,d\mu(x)=D(f).$$

最後に，

$$D(\varphi_t)=\int\Delta(\varphi_t(x))\,d\mu(x)$$
$$=\int\Delta(t)\,d\mu(x)=\Delta(t).\quad\blacksquare$$

この定理の逆も証明されるが，それらを合わせると，Riesz 型の定理が実現されることになり，fuzzy 集合論がいよいよ函数解析の手法で論じられることになる．本節ではそれらの詳細な定理の形と証明は省略するが，いずれ他の機会に詳しく論ずる予定である．

第5章　通信における数理解析

通信の理論を数学的に展開するには様々の手段が用いられるが，本章においては，確率論をベースとして話を展開する．確率的に話を進めるとき，多くは有限事象系において論じられている．しかし，多くの通信機構における現象を理論的に取り扱う場合は，有限の構造のみで把握することは不可能である．通常，何らかの有限個の記号集団を基として，有限の長さの記号の鎖の族を構成し，統計的データに基づいてその鎖に確率構造が付与され，数学的議論が展開されていく．実際に，我々が遭遇する事象として，10 円銅貨を投げて表裏でかけを行ったり，さいころを振って丁，半でかけを行ったりするなど，一定の有限回数で決まりを付けないで際限なく(不特定回数)試行が続行される場合，あるいは通信文をひとつのアルファベットを用いて長時間伝送するなど，あらかじめ定まった特定の有限の時間で終点に達しない場合を論ずるためには，無限の構造をもった観測系を導入する必要がある．本章では，メッセージ空間と情報源の話（5.1 節）から始め，McMillan の定理を証明し (5.2 節)，チャンネルの様々な特質について論ずる（5.3 節および 5.4 節）．さらに，情報伝送の中心的概念である相互エントロピーを 5.5 節で導入し，その主たる性質を調べる．最後に，5.6 節で符号化定理を概説する．

5.1　メッセージ空間と情報源

$A=\{x^{(1)},\cdots,x^{(a)}\}$ をひとつのアルファベットの集合（a は A の元の個数）とする．我々が A の要素を用いて文章を書いたり，言葉を発したり，信号を出したりする日常の行動は，遡った過去からはるかな未来に向かってとられる．そこで，単位時間ごとにいずれかの $x^{(k)}$ を A から取り出すとする．その出発時刻を s，終了時刻を t（$s\leq t$）とすると

$$\prod_{k=s}^{t} x_k = (x_s, x_{s+1}, \cdots, x_{t-1}, x_t) \tag{5.1}$$

というような文字の鎖(列)を得る．ここで x_k はすべて A の要素．これはアルファベット A を $t-s+1$ 回だけ掛け合わせた集合（直積集合）$A\times\cdots\times A=\prod_{k=s}^{t} A_k=A^{t-s+1}$ の要素と考えてよい．つまり，s, t は任意の整数で，$s\leq t$ という約束のみで記号 $\prod_{k=s}^{t} x_k$ が組み立てられる．したがって，まず始点，終点を設けないで

$$\cdots, x_s, x_{s+1}, \cdots, x_{s+n}, \cdots \quad (\text{これを } \Pi x_k \text{ とおく}) \tag{5.2}$$

のように無限列を考え，この無限列全体のなす集合は無限直積集合で

$$A^{\boldsymbol{Z}} = \prod_k A_k \quad (\text{ただし，} A_k \equiv A, \ \forall k \in \boldsymbol{Z}) \tag{5.3}$$

と表し，次に始点が s で，終点が t $(s\leq t)$ である文字の鎖を記号

$$[x_s{}^0 \cdots x_t{}^0] \equiv \{\Pi x_k \in A^{\boldsymbol{Z}};\ x_k = x_k{}^0 \ (s\leq k\leq t)\} \tag{5.4}$$

によって表すことにする．この鎖を**メッセージ**（message）と名付ける．これは $A^{\boldsymbol{Z}}$ の部分集合である．

メッセージは，必ずしも続けて出されるものではなく断続し，とぎれとぎれに出されたり，文字を出す何らかの（機械などの）都合により，ある間隔で複数の文字列を出すこともある．これらを集合演算で表示すると，前者は

$$x_i{}^0, \cdots, x_j{}^0;\ x_k{}^0, \cdots, x_l{}^0;\ x_m{}^0, \cdots, x_n{}^0 \quad (i\leq j<k\leq l<m\leq n)$$

で，これを

$$[x_i{}^0 \cdots x_j{}^0] \cap [x_k{}^0 \cdots x_l{}^0] \cap [x_m{}^0 \cdots x_n{}^0];$$

と表し，後者は

$$x_i{}^0, \cdots, x_s{}^0;\ y_i{}^0, \cdots, y_s{}^0;\ z_i{}^0, \cdots, z_s{}^0$$

で，これを

$$[x_i{}^0 \cdots x_s{}^0] \cup [y_i{}^0 \cdots y_s{}^0] \cup [z_i{}^0 \cdots z_s{}^0]$$

と表す．各々3個の $[\cdots]$ の場合を記したが，任意有限個数で行っても同じである．これらも総称して**メッセージ**とよぶ．メッセージは，簡単な集合演算によって次の関係式を満たすことが容易にわかる．

（1）　$[x_i{}^0 \cdots x_s{}^0] \subset [x_j{}^0 \cdots x_t{}^0]$ $(i\leq j\leq t\leq s)$;

（2）　$[x_i{}^0 \cdots x_s{}^0] \neq [y_i{}^0 \cdots y_s{}^0] \Longleftrightarrow [x_i{}^0 \cdots x_s{}^0] \cap [y_i{}^0 \cdots y_s{}^0] = \emptyset$;

(3) $[x_i{}^0\cdots x_s{}^0]=\cap\{[x_k{}^0];\ k=i,\cdots,s\}$;

(4) 任意の i,j,s,t $(i\leq j\leq s\leq t)$ に対して

$$[x_j{}^0\cdots x_s{}^0]=\cup\{[x_i\cdots x_t];\ x_k=x_k{}^0,\ \ k=j,\cdots,s\};$$

(5) $[x_j{}^0\cdots x_s{}^0]^c$(補集合)$=\cup\{[x_j\cdots x_s];\ x_k\neq x_k{}^0,$ ある $k=j,\cdots,s\}$;

(6) $\cup\{[x_i];\ x_i\in A\}=\bigcup\{[x_i\cdots x_s];\ x_i,\cdots,x_s\in A\}=A^Z.$

メッセージの全体を $\mathscr{M}$ で表すと，上記の計算によって $\mathscr{M}$ は集合体をなすことが知られる：$E,F\in\mathscr{M}\Longrightarrow E\cup F,\ E\cap F,\ E^c\in\mathscr{M}$. この $\mathscr{M}$ によって生成される σ-集合体を $\mathscr{F}_A$ で表すと，可測空間 $(A^Z,\mathscr{F}_A)$ を得る．これを**メッセージ空間**と名付ける．各メッセージを**筒集合** (cylinder set) とよぶこともある．以上が直積集合 A^Z の可測構造の話である．

ここで，応用上有効な A^Z の**位相構造**について述べよう，A は有限集合であるので，離散位相を考えてコンパクトであり，次のことがいえる：

1° 直積位相に関して，A^Z はコンパクト Hausdorff 空間となる．

直積位相は Tychonoff の積位相によって導入されるが，その定義によって

2° メッセージの族 $\mathscr{M}$ は位相基をなしている．位相基としては $\{[x_s\cdots x_t];\ s,t\in Z,s<t\}$ をとれば十分である．これは可算族をなす．

3° 位相空間 A^Z は第二可算公理を満たし，距離付け可能である．

4° A^Z の距離は通常，次の方法で定義される：まず，$r(x^{(i)},x^{(j)})=|i-j|/a$ $(\forall x^{(i)},x^{(j)}\in A)$ とおく．これを用いて，

$$d(x,y)=\sum_{n=-\infty}^{\infty}r(x_n,y_n)/2^{|n|}\quad(x=\Pi\,x_n,\ \ y=\Pi\,y_n\in A^Z)$$

とおくと，d は A^Z 上で距離函数の条件を満たし，この d で A^Z は上述の積位相と同相なコンパクト距離空間となる．

位相空間 A^Z においてすべての $E\in\mathscr{M}$ は開集合であり，また $E\in\mathscr{M}\Rightarrow E^c\in\mathscr{M}$ (関係式 (5)) から E^c も開集合である．したがってメッセージはすべて閉かつ開な集合である．すなわち，

5° $E\subset A^Z$ がメッセージ (i.e., $E\in\mathscr{M}$) $\Longleftrightarrow E$ は閉かつ開な集合．

このように閉かつ開な集合の族が位相基である位相空間を**完全不連結**という．

写像 $T:\Pi x_k \in A^Z \to \Pi x_k' \in A^Z, x_k'=x_{k+1}$ $(k=0,\pm1,\pm2,\cdots)$ を**シフト変換**という．これに関して

6°　この T は，$T^{-1}(\Pi x_k)=\Pi x_{k-1}$ を満たし，A^Z から A^Z への全単射な双連続写像である．

7°　$T^{-n}[x_s{}^0\cdots x_t{}^0]=[y_{s+n}{}^0\cdots y_{t+n}{}^0]$，$x_k{}^0=y_{k+n}{}^0$ $(s\leq k\leq t)$．

コンパクト距離空間 A^Z 上の複素数値連続函数の全体を $C(A^Z)$ とする．$C(A^Z)$ は線形空間であり

8°　任意の $f\in C(A^Z)$ に対し $|f|$ は最大値をもつ：$\|f\|\equiv\sup\{|f(x)|; x\in A^Z\}=\max\{|f(x)|; x\in A^Z\}$．この $\|\cdot\|$ はノルムの条件を満たし，$C(A^Z)$ は可換な C^* 代数となる（i.e.，$\|f^*f\|=\|f\|^2$ をみたす Banach* 代数（p. 99））．

9°　$f^T(x)=f(Tx)$ とおくと $f\to f^T$ は＊自己同型，i.e.，$(\alpha f+\beta g)^T=\alpha f^T+\beta g^T$，$(fg)^T=f^Tg^T$，$f^{*T}=f^{T*}$，$\{f^T; f\in C(A^Z)\}=C(A^Z)$．変換 $f\to f^T$ を作用素として扱う場合，$f^T=f\circ T$ または Tf と表すこともある．

C^* 代数と対合＊に関する議論は第2巻で詳しく論じてある．これらの概念に不案内な読者はこの章では飛して読んでもかまわない．

ここまでは，メッセージ空間 A^Z の数学的構造について述べたが，A^Z における各メッセージ $[x_s\cdots x_t]$ はでたらめに出現するものではない．メッセージを発する人，物，機械など特定の性格，機能をもった“もの”によってつくり出される．時刻 k で A からどの要素が選ばれるかは，その“もの”によって定まる．しかも，多くは決定論的に定まるのではなく，誤差とか思い違いなどによったり，あるいは何らかの偶然性に支配されて選出される場合もある．つまり，その特定の“もの”のもつ特徴が，付随した確率分布に従うという数学的基盤に立って把握される．

まず，簡単な場合から論じよう．アルファベット A が与えられ，それを扱う“もの”が定まると，完全事象系 $\begin{pmatrix}A\\P\end{pmatrix}$ が設定される．すると，メッセージの確率が

$$\mu([x_s{}^0\cdots x_t{}^0])=\prod_{k=s}^{t}p(x_k{}^0)$$

で定まる．集合演算を併せ行うことにより，p はメッセージからなる集合代数

$\mathcal{M}$ 上の加法的な $[0,1]$-値函数として一意に定義され，$A^Z=\bigcup_i [x_k{}^i]$, $[x_k{}^i]\cap [x_k{}^j]=\emptyset$ $(i\neq j)$ であるから

$$\mu(A^Z)=\mu(\bigcup_i [x_k{}^i])=\sum_i \mu([x_k{}^i])=\sum_i p(x_k{}^i)=1.$$

ここで Kolmogorov の拡張定理を適用し，加法的集合函数 $\mu:\mathcal{M}\to[0,1]$ が σ-加法的集合函数 $\mathcal{F}_A\to[0,1]$ に一意に拡張される．これも同一の記号 μ で表すと，μ は確率測度となり，確率空間 $(A^Z,\mathcal{F}_A,\mu)$ が定義される．この確率測度 μ を **Bernoulli 型**という．

上述の場合は，メッセージの出現確率がアルファベット A に付随した確率分布に従う場合であったが，これは特別な場合であって，一般には各メッセージそのものに出現確率が対応している場合を論ずる必要がある．つまり，各 $E\in\mathcal{F}_A$ に確率 $0\leq\mu(E)\leq 1$ が対応する場合である．この確率空間 $(A^Z,\mathcal{F}_A,\mu)$ を**情報源**と名付ける．これを以下 $[A^Z,\mu]$ と略記する．とくに μ が **Bernoulli 型**のとき，$[A^Z,\mu]$ を **Bernoulli 情報源**という．

（1） μ がシフト変換 T に関して $\mu(T^{-1}E)=\mu(E)$, $E\in\mathcal{F}_A$ であるとき，μ は ***T*-定常**（または ***T*-不変**）であるという．このとき，$[A^Z,\mu]$ を**定常情報源**という．

（2） $E\in\mathcal{F}_A$ が $\boldsymbol{\mu}$-a.e. T-不変（すなわち，$\mu(T^{-1}E\triangle E)=0$. ここで，$C\triangle D=(C\cup D)\cap(C\cap D)^c$ で，これを対称差という）ならば，つねに $\mu(E)=0$ または 1 であるとき，μ を**エルゴード的**，$[A^Z,\mu]$ を**エルゴード情報源**という．Bernoulli 情報源は定常かつエルゴード情報源である．

（3） 情報源 $[A^Z,\mu]$ の測度 μ 全体（つまり $(A^Z,\mathcal{F}_A)$ 上の確率測度の全体）を $\boldsymbol{P}(A^Z)$, T-定常な $\mu\in\boldsymbol{P}(A^Z)$ の全体を $\boldsymbol{P}_T(A^Z)$, エルゴード的な $\mu\in\boldsymbol{P}_T(A^Z)$ 全体を $\boldsymbol{P}_E(A^Z)$ で表す．

これらの記号のもとで，以下函数解析に基づいた手法によって議論を進めよう．なお $C(A^Z)^*$ は $C(A^Z)$ 上の有界線形汎函数の全体とする．

10° $S_1\equiv\{\xi\in C(A^Z)^*\,;\,\xi(1)=\|\xi\|=1\}$ とおくと，$\xi\in S_1\Longleftrightarrow\xi(f)\geq 0$ $(\forall f\in C(A^Z), f\geq 0)$ かつ $\|\xi\|=1$. ここで，各 $\mu\in\boldsymbol{P}(A^Z)$ に対して

$$\dot{\mu}(f)=\int_{A^Z} f(x)d\mu(x),\qquad \forall f\in C(A^Z),$$

とおくと $S_1=\{\dot{\mu};\mu\in \boldsymbol{P}(A^Z)\}$. この事実によって μ と $\dot{\mu}$ を同一視することができ，$\boldsymbol{P}(A^Z)=S_1$ と考えることができる（このことは Riesz の表現定理のひとつの場合である）.

11°　$\mu\to\mu\circ T^{-1}$ はコンパクトな凸集合 $\boldsymbol{P}(A^Z)$ 上のアフィン変換である. したがって，Kakutani-Markov の不動点定理（付章5節参照）より，$\boldsymbol{P}_T(A^Z)\neq\emptyset$.

12°　$\boldsymbol{P}(A^Z)$, $\boldsymbol{P}_T(A^Z)$ は $C(A^Z)^*$ の弱*位相に関してコンパクトであり，ともに凸集合であるから，Krein-Milman の定理によってともに十分多くの端点をもち

$$\mathrm{ex}\ \boldsymbol{P}(A^Z)=\{\delta_x;\ x\in A^Z\},\qquad \mathrm{ex}\ \boldsymbol{P}_T(A^Z)=\boldsymbol{P}_E(A^Z)$$

が成立する. ここで δ_x は一点集合を台にもつ点測度であり，ex$K$（K は凸集合）は K の端点の全体を表す.

註　11°, 12° の命題は A^Z という特殊な空間でなく，一般のコンパクト Hausdorff 空間の場合にも成立する. さらに，このことを可換な C^* 代数に適応すると，Gelfand 表現定理が証明される（第2巻の付章2節2.2項).

13°　$\mu\in\boldsymbol{P}_T(A^Z)$ のとき，

（1） **von Neumann 平均エルゴード定理**：$f\in L^2(A^Z,\mu)$ に対して，$f_T\in L^2(A^Z,\mu)$ が μ-a.e. で一意に存在して，

$$f_T\circ T=f_T\quad かつ\quad \lim_{n\to\infty}\int_{A^Z}\left|\frac{1}{n}\sum_{k=0}^{n-1}f(T^kx)-f_T(x)\right|^2 d\mu=0.$$

（2） **Birhkoff 個別エルゴード定理**：$f\in L^1(A^Z,\mu)$ に対して，$f_T\in L^1(A^Z,\mu)$ が μ-a.e. で一意に存在して，

$$f_T\circ T=f_T\quad かつ\quad \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}f(T^kx)=f_T(x),\ \ \mu\text{-a.e.}.$$

ここで,（1）は $f\in L^2$ の時間平均が L^2-収束することであるが，$f\in L^p(p\geqq 1)$ の場合にも同様に L^p-収束が成立する［H.3, T.4, Y.3].

14°　上の（1),（2）において，$f=1_{[x_s\cdots x_t]}$ とおくと，$\mu\in\boldsymbol{P}_T(A^Z)$ がエルゴード的（i.e., $\mu\in\boldsymbol{P}_E(A^Z)$）であるための必要十分条件は，すべてのメッセ

ージについて，時間平均と空間平均が一致することである：

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}1_{[x_s\cdots x_t]}(T^k x)=\mu([x_s\cdots x_t]).$$

上述の 13° において，

$$f_n(x)=\frac{1}{n}\sum_{k=0}^{n-1}f(T^k x)$$

とおくことによって 13° の極限の式（1）と（2）は簡単に記述できる．この記号を用いて

15° $\mu\in \boldsymbol{P}_T(A^{\boldsymbol{Z}})$ に対して次の条件は同値である：

（1） $\mu\in \boldsymbol{P}_E(A^{\boldsymbol{Z}})=\mathrm{ex}\,\boldsymbol{P}_T(A^{\boldsymbol{Z}})$.

（2） $\lim_{n\to\infty}\mu(f_n g)=\mu(f)\mu(g),\quad \forall f,g\in C(A^{\boldsymbol{Z}})$.

（3） $\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\mu(E\cap T^{-k}F)=\mu(E)\mu(F),\quad \forall E,F\in\mathcal{F}_A$.

註　13°，15° の命題は一般の確率測度空間 $(X,\mathcal{F}_X,\mu)$ で，保測変換 T をもつ場合においても成立する，ただし 15° の（2）の $C(A^{\boldsymbol{Z}})$ は有界 $\mathcal{F}_X$-可測函数の全体 $\boldsymbol{B}(X)$ に置き換える．

この節の議論は主に文献 [B.7, K.12, U.9] からとっている．

5.2 情報源のエントロピーと McMillan の定理

本節では，簡単のため定常情報源 $[A^{\boldsymbol{Z}},\mu]$ について考える．$A^{\boldsymbol{Z}}$ の有限可測分割を $\tilde{C},\tilde{D}$ 等で表そう．ここで，第2章の 2.2 節を振り返ることにする．この $\tilde{C},\tilde{D}$ に対して，これらの合成によって第三の分割 $\tilde{C}\vee\tilde{D}$ が得られ，またシフト T を作用させて分割 $T^{-1}\tilde{C}$ を得る．さらに，条件付エントロピー

$$S(\tilde{C}|\tilde{D})=-\sum_{C\in\tilde{C},D\in\tilde{D}}\left\{\frac{\mu(C\cap D)}{\mu(D)}\log\frac{\mu(C\cap D)}{\mu(D)}\right\}\mu(D)\qquad(5.5)$$

が定義される．これは，$\tilde{D}$ が既知という条件下で分割 $\tilde{C}$ の観測によって得る情報量を意味している．このとき，

（1） $S(\tilde{C}\vee\tilde{D})=S(\tilde{C}|\tilde{D})+S(\tilde{D})=S(\tilde{D}|\tilde{C})+S(\tilde{C})$.

（2） $S(\tilde{C})\leq S(\tilde{C}\vee\tilde{D})$.

（3） $S(\tilde{C}|\tilde{D})\leq S(\tilde{C})$.

（4）　$S(\tilde{C}\vee\tilde{D})\leq S(\tilde{C})+S(\tilde{D})$.

が成立する（〈定理 2.1〉の単なる書き換えである）.

ここで，メッセージ空間 A^Z の固有の性質が直接反映する事柄について議論する．$\mathscr{M}_n$ を長さ n のメッセージの族

$$\mathscr{M}_n=\{[x_0{}^{(1)},\cdots,x_{n-1}{}^{(1)}],\ [x_0{}^{(2)},\cdots,x_{n-1}{}^{(2)}],\cdots\}\ (=\bigvee_{k=0}^{n-1}T^{-k}\mathscr{M}_1\ \text{である})$$

とする．この $\mathscr{M}_n$ は，A^Z の有限可測分割である．さらに，$\mathscr{M}_n$ の各要素は閉かつ開であり，アルファベット A の要素の個数が a のとき，$\mathscr{M}_n$ の要素の個数は a^n である．したがって，$\mathscr{M}_n$ のエントロピーは

$$S(\mathscr{M}_n)=-\sum_{k=1}^{a^n}\mu([x_0{}^{(k)}\cdots x_{n-1}{}^{(k)}])\log\mu([x_0{}^{(k)}\cdots x_{n-1}{}^{(k)}])$$

となり，これはメッセージの始時点（$\mathscr{M}_n$ では0とした）がどんな値をとるかということには無関係であり，情報源の分布 μ とメッセージの長さ n によって一意に定まる．すなわち，情報量の定義から，情報源 $[A^Z,\mu]$ から出される長さ n のメッセージは平均して $S(\mathscr{M}_n)/n$ の情報量をもっていることがわかる．次に，これの極限についての重要な定理を証明する．

〈定理 5.1〉　$0\leqq S(\mu)\equiv\lim_{n\to\infty}S(\mathscr{M}_n)/n\leqq\log a$.

この極限 $S(\mu)$ はメッセージ空間 A^Z を定常情報源にならしめる測度 μ を設置したとき，情報源の各メッセージの1文字当たりの平均情報量を表しており，これを測度 μ で表せる**情報源エントロピー**とよぶ．この章では T を固定し μ を動かすので，$S(\mathscr{M}_1,T)$ と書く（第2章）代わりに，$S(\mu)$ と書く．

証明　任意の $m,n=1,2,\cdots$ に対して，

$$S_m\equiv S(\mathscr{M}_m),\quad \mathscr{M}_n\equiv\{[y_m{}^{(1)}\cdots y_{m+n-1}{}^{(1)}],[y_m{}^{(2)}\cdots y_{m+n-1}{}^{(2)}],\cdots\}$$

とおくと，$\mathscr{M}_{m+n}=\mathscr{M}_m\vee\mathscr{M}_n$ となり，（2）と（4）によって

$$S_m\leqq S_{m+n}\leqq S_m+S_n,$$

$$S_n\leqq S_{2n}\leqq 2S_n,\qquad S_{kn}\leqq kS_n\quad(k=1,2,\cdots;\ n=1,2,\cdots).$$

とくに $n=1$ として，$S_k\leqq kS_1$，すなわち，$S_k/k\leqq S_1$ $(k=1,2,\cdots)$．ゆえに，

$$c \equiv \limsup_{k\to\infty} \frac{S_k}{k} < +\infty \tag{5.6}$$

が存在する．これから，任意の $\varepsilon>0$ に対して，自然数 k を適当にとると，$S_k/k < c+\varepsilon$．また任意の $n>k$ に対して

$$m>1 \quad \text{かつ} \quad k(m-1)<n\leq km$$

のように m をとる（これはつねに可能）．このとき $S_n \leq S_{mk}$ であり，

$$\frac{S_n}{n} \leq \frac{S_{km}}{n} \leq \frac{S_{km}}{k(m-1)} \leq \frac{m}{m-1}\frac{S_k}{k}.$$

ここで，n を十分大にとれば m も十分大になり，極限の式 (5.6) により，

$$c-\varepsilon \leqq S_n/n < c+2\varepsilon \quad \text{ゆえに} \quad c=\lim_{n\to\infty} S_n/n.$$

なお，$\mathscr{M}_1=\{[x_0^{(1)}],[x_0^{(2)}],\cdots,[x_0^{(a)}]\}$ とかけるから，

$$\frac{1}{n}S_n \leqq S_1 = -\sum_{k=1}^{a} \mu[x_0^{(k)}]\log\mu[x_0^{(k)}] \leqq \log a \quad (\forall n). \ \blacksquare$$

なお，この定理は，μ の T-不変性を取りはずしても，$a\equiv|A|$ が有限である限り成立することに注意しよう．

次に，第2章と同様 $\widetilde{C}, \widetilde{D}, \widetilde{E}_1, \widetilde{E}_2, \cdots, \widetilde{E}_n$ 等を A^Z の有限可測分割とする．これらに対して第2章のように，エントロピー函数と条件付エントロピー函数は各々次で定める：

$$I(\widetilde{C})(x) = -\sum_{C\in\widetilde{C}} \log\mu(C)1_C(x),$$

$$I(\widetilde{C}|\widetilde{D})(x) = -\sum_{C\in\widetilde{C}} \log P(C|\widetilde{D})1_C(x).$$

ここで，$P(C|\widetilde{D})$ は有限分割 $\widetilde{D}$ の生成する有限集合体に関する C の条件付確率であり，条件付期待値も同様に定義する（有限集合体に対しても，そのアトム全体をとることによって一意に有限分割が定まり，よって有限分割と有限集合体は1対1に対応する）．このとき明らかに，

$$I(\widetilde{C}) = I(\widetilde{C}|\mathcal{O}), \qquad I(\widetilde{C}\vee\widetilde{D}) = I(\widetilde{C}) + I(\widetilde{C}|\widetilde{D}). \tag{5.7}$$

ここで $\mathcal{O}$ は null 集合と全空間 A^Z からなる集合体を表す．さらに，

$$I(\widetilde{C}|\widetilde{D})\circ T = I(T^{-1}\widetilde{C}|T^{-1}\widetilde{D}), \qquad I(\widetilde{C})\circ T = I(T^{-1}\widetilde{C}) \tag{5.8}$$

が成立した（第2章参照）．ここで，以下の議論のために，いくつかの命題を

証明しておこう.

1°　有限分割の列 $\widetilde{E}_0=\mathcal{O}, \widetilde{E}_1, \cdots, \widetilde{E}_n$ $(n=1,2,\cdots)$ に対して

$$I\left(\bigvee_{i=1}^{n}\widetilde{E}_i\right)=\sum_{k=1}^{n}I\left(\widetilde{E}_k\,\middle|\,\bigvee_{i=0}^{k-1}\widetilde{E}_i\right)$$

証明　$n=1$ のときは, (5.7) によって $I(\widetilde{E}_1)=I(\widetilde{E}_1|\mathcal{O})$. 一般の $n>1$ に対しては, 同じく (5.7) により

$$I\left(\bigvee_{i=1}^{n+1}\widetilde{E}_i\right)=I\left(\left(\bigvee_{i=1}^{n}\widetilde{E}_i\right)\vee\widetilde{E}_{n+1}\right)$$

$$=I\left(\bigvee_{i=1}^{n}\widetilde{E}_i\right)+I\left(\widetilde{E}_{n+1}\,\middle|\,\bigvee_{i=1}^{n}\widetilde{E}_i\right).$$

以下, 帰納的に $I\left(\bigvee_{i=1}^{n}\widetilde{E}_i\right)$, $I\left(\bigvee_{i=1}^{n-1}\widetilde{E}_i\right),\cdots$ に対して各々同様な分割を行うことによって結論を得る. ▌

2°　ひとつの分割 $\widetilde{E}$ に対して

$$I\left(\bigvee_{i=0}^{n-1}T^{-i}\widetilde{E}\right)=I(T^{-(n-1)}\widetilde{E})+\sum_{k=1}^{n-1}I\left(T^{-(n-k-1)}\widetilde{E}\,\middle|\,\bigvee_{i=1}^{k}T^{-(n-i)}\widetilde{E}\right)$$

これは, 1° において $\widetilde{E}_i=T^{i-n}\widetilde{E}$ $(i=1,2,\cdots,n)$ とおけばすぐいえる. この結果に等式 (5.8) を適用すると, ただちに

3°　$I\left(\bigvee_{i=1}^{n-1}T^{-1}\widetilde{E}\right)=I(\widetilde{E})\circ T^{n-1}+\sum_{k=1}^{n-1}I\left(\widetilde{E}\,\middle|\,\bigvee_{i=1}^{k}T^{-i}\widetilde{E}\right)\circ T^{n-k-1},\quad n>1.$

ここでさらに,

$$\left.\begin{aligned}&f_n\equiv I\left(\bigvee_{k=0}^{n-1}T^{-k}\widetilde{C}\right),\\&g_0\equiv I(\widetilde{C}),\qquad g_n\equiv I\left(\widetilde{C}\,\middle|\,\bigvee_{k=1}^{n}T^{-k}\widetilde{C}\right),\ g\equiv I\left(\widetilde{C}\,\middle|\,\bigvee_{k=1}^{\infty}T^{-k}\widetilde{C}\right)\end{aligned}\right\}\quad n=1,2,\cdots \tag{5.9}$$

とおくと

4°　$f_n=\sum_{k=0}^{n-1}g_k\circ T^{n-k-1},\qquad n=1,2,\cdots$

これは 3° の等式のかき換えである ($\widetilde{E}=\widetilde{C}$ とおき直せばよい).

5°　$\{g_n\}$ は一様可積分である $\left(\text{i.e.,}\ \sup_n\int_{|g_n|>\lambda}|g_n|d\mu\to 0\ (\lambda\to\infty)\right)$.

証明　ひとつの $C\in\widetilde{C}$ に対して,

$$\widetilde{D}_n=\bigvee_{k=1}^{n} T^{-k}\widetilde{C}, \qquad p_n(x)=-\log P(C|\widetilde{D}_n)(x), \qquad q_n(x)=p_n(x)1_C(x)$$

とおくとき，$\{q_n\}$ の一様可積分性をいえばよい．$0<\alpha<\beta<\infty$ に対して

$$E\equiv\{x\in A^Z;\ \alpha\leqq p_n(x)\leqq\beta\}\in\sigma(\widetilde{D}_n).$$

ここで $\sigma(\widetilde{D})$ は $\widetilde{D}$ が生成する σ 集合体を表す．したがって

$$\log P(C|\widetilde{D}_n)(x)\leqq-\alpha, \qquad \text{i.e.}, P(C|\widetilde{D}_n)(x)\leqq e^{-\alpha}(\forall x\in E).$$

さらに

$$F_{\alpha,\beta}=\{x\in A^Z;\ \alpha\leqq q_n(x)\leqq\beta\}$$

とおくと，$C\cap E=F_{\alpha,\beta}$ であり

$$\mu(F_{\alpha,\beta})=\mu(C\cap E)=\int_E 1_C d\mu$$

$$=\int_E P(C|\widetilde{D}_n)d\mu\leqq e^{-\alpha}\mu(E)\leqq e^{-\alpha}.$$

一方，

$$\int_{F_{\alpha,\beta}\cap C} q_n d\mu=\int_{F_{\alpha,\beta}} q_n d\mu\leqq\beta\mu(F_{\alpha,\beta})\leqq\beta e^{-\alpha}. \tag{5.10}$$

ここで，$\alpha=n$，$\beta=n+1$ とすると

$$\{q_n\geqq\lambda\}=\bigcup_{k=0}^{\infty}\{k+\lambda\leqq q_n\leqq k+\lambda+1\}=\bigcup_{k=0}^{\infty}F_{\lambda+k,\lambda+k+1}.$$

(5.10) より，

$$\int_{\{q_n\geqq\lambda\}} q_n d\mu=\sum_{k=0}^{\infty}\int_{F_{\lambda+k,\lambda+k+1}} q_n d\mu$$

$$\leqq\sum_{k=0}^{\infty}(\lambda+k+1)e^{-(\lambda+k)}$$

$$=e^{-\lambda}\sum_{k=0}^{\infty}(\lambda+k+1)e^{-k}$$

$$\to 0 \quad (\lambda\to\infty).$$

この収束は n に関して一様である．▌

6° $g_n\to g$ a.e. かつ L^1-収束.

証明 各 $C\in\widetilde{C}$ に対して $P(C|\widetilde{D}_n)$ は martingale である．$\widetilde{D}_\infty\equiv\bigvee_{k=1}^{\infty}T^{-k}\widetilde{C}$ とおくと，

$$P(C|\tilde{D}_n)\to P(C|\tilde{D}_\infty)\ \text{a.e.}\quad \text{かつ}\quad L^1\text{-収束}$$

が成立する（第2章，2.1節の21°）．したがって

$$q_n=1_C\log P(C|\tilde{D}_n)\to q=1_C\log P(C|\tilde{D}_\infty)\ \text{a.e.}.$$

5°によって $\{q_n\}$ は一様可積分であるから，L^1-ノルムに関して一様有界性をいえば，これが L^1-ノルム収束することが結論できる．これを示す．$q_n\geqq 0$ であり，各 $\log P(C|\tilde{D}_n)$ は台の上で有限値函数であるから，

$$\int q_n\,d\mu=-\int 1_C\log P(C|\tilde{D}_n)\,d\mu$$

$$=-\int P(C|\tilde{D}_n)\log P(C|\tilde{D}_n)\,d\mu\leqq 1,\qquad n=1,2,\cdots.\ \blacksquare$$

7° $h\in L^1(A^Z,\mu)$ が存在して $h\circ T=h$，かつ

$$\frac{1}{n}\sum_{k=0}^{n-1}g\circ T^{n-k-1}\to h\quad \text{a.e.}\quad \text{かつ}\quad L^1\text{-収束}.$$

これは前述の個別エルゴード定理と平均エルゴード定理そのものである．

ここでいよいよ目標の定理に到達する．これまで使った記号 f_n, g_n, g, h （(5.9) 参照）を用いると，

〈**定理 5.2**〉 **McMillan の定理** 函数列 $\{f_n\}$ は T-不変な函数 h に a.e. かつ L^1-収束する：

$$\frac{1}{n}f_n\left(\equiv\frac{1}{n}I\left(\bigvee_{k=0}^{n-1}T^{-k}\tilde{C}\right)\right)\to h.$$

とくに $[A^Z,\mu]$ がエルゴード情報源のとき，

$$h=S(\tilde{C},T)=\int g\,d\mu\,(=T\ \text{と}\ \tilde{C}\ \text{に関する}\ K\text{-}S\ \text{エントロピー})\ \text{a.e.}$$

証明 L^1-収束を証明する．

$$\left\|\frac{1}{n}f_n-h\right\|_1=\left\|\frac{1}{n}\sum_{k=0}^{n-1}g_k\circ T^{n-k-1}-h\right\|_1$$

$$\leqq\left\|\frac{1}{n}\sum_{k=0}^{n-1}(g_k\circ T^{n-k-1}-g\circ T^{n-k-1})\right\|_1+\left\|\frac{1}{n}\sum_{k=0}^{n-1}g\circ T^{n-k-1}-h\right\|_1$$

$$\leqq\frac{1}{n}\sum_{k=0}^{n-1}\|g_k-g\|_1+\left\|\frac{1}{n}\sum_{k=0}^{n-1}g\circ T^{k}-h\right\|_1.$$

これは 7° より，$n\to\infty$ のとき0に収束する．

a.e. 収束は，ほぼ同様の仕方で，6°, 7° の結果と〈補題 2.8〉と〈補題2.9〉を使って証明することができる．

定理の後半は定義より明らか．▮

註　〈定理 5.2〉は定常情報源 $[A^{\mathbf{Z}},\mu]=(A^{\mathbf{Z}},\mathcal{F}_A,\mu,T)$ で一般な有限分割 $\tilde{C}$ を固定して，T によるシフトを繰り返し，極限を計算することによって証明された．ここで $\mathcal{F}_A$ は $A^{\mathbf{Z}}$ の Borel 集合体，T はシフト変換で μ を定常（$=T$-不変）とした．その証明法を何ら変更することなしに〈定理 5.2〉は任意の確率測度空間 $(X,\mathcal{F}_X,\mu,T)$ において成立する．ここで T は X 上の保測変換である．いずれにしても，極限の函数 h は最初に与えられた可測分割 $\tilde{C}$ に従属して定まる．この一般論は，数学として意味ある一種の平均エルゴード定理と見なせる．

さて，〈定理 5.2〉は情報理論として重要な定理であるが，それを説明するためには，やはり情報源 $[A^{Z},\mu]$ において，一般の可測分割ではなくて，メッセージによる分割の列を採用して論ずる必要がある．72 ページで与えた記号を用いて，まずメッセージによる分割の列

$$\mathcal{M}_1,\mathcal{M}_2,\mathcal{M}_3,\cdots$$

を考えよう．各々にシフト変換を作用させると，等式

$$\mathcal{M}_n=\mathcal{M}_{n-1}\vee T^{-1}\mathcal{M}_{n-1}=\bigvee_{k=0}^{n-1}T^{-k}\mathcal{M}_1$$

が成立する．〈定理 5.2〉においては，分割 $\tilde{C}$ をひとつ定め，

$$f_n=I\left(\bigvee_{k=0}^{n-1}T^{-k}\tilde{C}\right)=-\sum_{E\in\bar{E}}1_E\log\mu(E),$$

$\bar{E}=\bigvee_{k=0}^{n}T^{-k}\tilde{C}$ としたのであるが，$\tilde{C}$ としてどの分割 $\mathcal{M}_1,\mathcal{M}_2,\cdots$ を採用しても結果は同じ f_n（n はそのとり方に対してある番号より先にする）を得る．したがって，

$$f_n=-\sum 1_{[x_0\cdots x_{n-1}]}\log\mu([x_0\cdots x_{n-1}])$$
$$(\textstyle\sum \text{ は } [x_0\cdots x_{n-1}]\in\mathcal{M}_n \text{ の和})$$

と表すことができ，〈定理 5.2〉は次の形にかかれる．

〈定理 5.3〉　定常情報源 $[A^{Z},\mu]$ において，シフト T に関して不変な函数 $h\in L^1(A^{Z})$ が存在して

$$-\frac{1}{n}\sum 1_{[x_0\cdots x_{n-1}]}\log\mu([x_0\cdots x_{n-1}])\to h \quad (\text{a.e. かつ } L^1\text{-収束})$$

が成立する．もし $[A^Z\mu]$ がエルゴード情報源ならば

$$h=S(\mu)\,(=\text{情報源のエントロピー})\ \text{a.e.}$$

である．

このMcMillanの定理は次のことを意味している．エルゴード情報源 $[A^Z\mu]$ においては，$\varepsilon>0$, $\delta>0$ がいかに小さくても，十分大きな n をとれば

$$\mu\left(\left\{x\in A^Z;\ \left|\frac{1}{n}f_n(x)-S(\mu)\right|>\varepsilon\right\}\right)<\delta \tag{5.11}$$

となる．したがって，このとき与えられた情報源から出てくる十分大きな n に対するメッセージの全体 $\mathcal{M}_n$ は，(5.11) により，2つの組 $\mathcal{M}_n^{(1)}$, $\mathcal{M}_n^{(2)}$ に分けられる：$M\in\mathcal{M}_n^{(1)}$ に対しては，

$$\left|\frac{1}{n}\log\mu(M)+S(\mu)\right|<\varepsilon$$

が成立し，$M\in\mathcal{M}_n^{(2)}$ に対しては，

$$\mu(\textstyle\bigcup_{M\in\mathcal{M}_n^{(2)}}M)<\delta$$

が成立する．すなわち，すべての $M\in\mathcal{M}_n^{(1)}$ の生起確率 $\mu(M)$ は大体等確率 e^{-nS} に近いことがわかり，このことから，$\mathcal{M}_n^{(1)}$ の要素の個数はほぼ e^{nS} に等しいことがわかる（ただし，$S=S(\mu)$）．$\mathcal{M}_n$ の要素の個数は $a^n=e^{n\log a}$ であるから，$S(\mu)\leqq\log a$ より，$S(\mu)=\log a$ の場合を除いて（これは通常起こらない．ほとんど $S(\mu)\ll\log a$），n が十分大きければ，$\mathcal{M}_n^{(1)}$ の要素の個数と比べて $\mathcal{M}_n^{(2)}$ の要素の個数の方がはるかに多いことがわかる．

このようにして，情報源のエントロピー $S(\mu)$ の構成とともにMcMillanの定理は，メッセージを十分長くとれば，そのメッセージの1語当たりの情報量が，情報源のそれに等しくするようにできること，すなわち，発信側の情報を再現する仕方が存在することを示している．よって，この定理は，前にも述べたように数学的にも評価されるが，情報理論としての価値が高いものである．

詳しくは，文献［H.4, K.12, M.2, P.1, U.14］などを参照．

5.3 チャンネルと平均作用素

2つのアルファベットの集合 A, B に対して，2つのメッセージ空間をそれぞれ $X=A^{\mathbf{Z}}$, $Y=B^{\mathbf{Z}}$ とおき，その積空間を $\Omega=X\times Y$ とおく．Ω と $(A\times B)^{\mathbf{Z}}$ は同一視することができ，Ω もメッセージ空間と考えられる．したがって X, Y, Ω はすべてコンパクト距離空間となり，各々の Borel 集合体を $\mathcal{F}_X, \mathcal{F}_Y, \mathcal{F}_\Omega$ で表す．また $(X, \mathcal{F}_X)$ 上の確率測度の全体を $\boldsymbol{P}(X)$，有界な Borel 可測函数の全体を $B(X)$ で表し，Y, Ω に対しても同様に $\boldsymbol{P}(Y)$, $\boldsymbol{P}(\Omega)$, $B(Y)$, $B(\Omega)$ を定義する．

註 5.1 節における記号と上記の記号は重複しているものがある，たとえば，

$$(X, \mathcal{F}_X)=(A^{\mathbf{Z}}, \mathcal{F}_A), \quad \boldsymbol{P}(X)=\boldsymbol{P}(A^{\mathbf{Z}}).$$

これらは，X, Y, Ω がメッセージ空間でなくとも，一般のコンパクト Hausdorff 空間としても同様である．X, Y 等が距離空間の場合は，Borel 可測＝Baire 可測であるので，したがって現在対象としている $X=A^{\mathbf{Z}}$, $Y=B^{\mathbf{Z}}$ 等では，$B(X)$, $B(Y)$, $B(\Omega)$ を有界 Baire 可測函数の全体と考えてよく，

1° $B(X)$ は $C(X)$ を含み有界単調極限について閉である (i.e., $\{f_n\}\subset B(X)$, $f_n\uparrow f$, $\|f_n\|\leq M\Rightarrow f\in B(X)$) 最小の函数族である．ここで $\|\cdot\|$ は sup-ノルムである．

2° $B(X)$ 上の有界線形汎函数で $\xi(1)=\|\xi\|=1$ (このときつねに $\xi(f)\geq 0$ $(\forall f\in B(X),\ f\geqq 0)$) かつ $\xi(f_n)\uparrow\xi(f)$ $(\forall\{f_n\},\ f\in B(X),\ f_n\uparrow f)$ である $\xi\in B(X)^*$ と $\mu\in\boldsymbol{P}(X)$ は，

$$\xi(f)=\int_X f(x)d\mu(x), \qquad f\in B(X)$$

によって1対1に対応している．したがって，$\boldsymbol{P}(X)$ と上記の ξ の全体は同一視できる．

3° 上の議論で $B(X)$ は $C(X)$ と同様の数学構造で可換 C^* 代数となり，$C(X)$ を C^* 部分代数として含むことがすぐわかる (実は，この $B(X)$ は Σ^* 代数といわれる代数になっている)．$C(X)$ の場合は $C(X)$ の状態全体の集合が $\boldsymbol{P}(X)$ と同一視できたが，$B(X)$ 上の状態 ξ とその $C(X)$ 上への制限

$\xi_0=\xi\upharpoonright C(X)$ は1対1に対応しない．それは，ξ に単調連続の条件を付けたものの，全体がまさに1対1に対応しているのである．つまり $C(X)$ のひとつの状態 ξ_0 をとり，それを $B(X)$ 上に Hahn-Banach 拡張を施して得られるものはすべて $B(X)$ の状態であるが，その中に唯ひとつ単調連続な ξ を見出せるということなのである．

2変数の函数 $\lambda(\cdot,\cdot):X\times\mathcal{F}_Y\to[0,1]$ ((0,1) 閉区間) が次の2つの条件

(C.1)　各 $x\in X$ に対して $\lambda(x,\cdot)\in\boldsymbol{P}(Y)$,

(C.2)　各 $F\in\mathcal{F}_Y$ に対して $\lambda(\cdot,F)\in B(X)$,

を満足するとき，3つの組 $[X,\lambda,Y]$ または単に λ を**チャンネル** (channel) といい，X を**入力源** (input)，Y を**出力源** (output) とよぶ．今，X,Y 上のシフト変換を同一の記号 T で表し，Ω 上のシフト変換も同じく $T(x,y)=(Tx,Ty)$ ($(x,y)\in X\times Y=\Omega$) によって T で表すことにする．こうした記法を用いることによっても，各々，定義域の変数を付して式を展開するため，混乱は生じない．5.1 節で導入された $\boldsymbol{P}_T(A^Z)$，$\boldsymbol{P}_E(A^Z)$ などは A^Z の代わりに X,Y,Ω を用いても全く同様である．とくに $\mu\in\boldsymbol{P}(\Omega)$ に対して確率空間 $[\Omega,\mu](=(\Omega,\mathcal{F}_\Omega,\mu))$ を**複合情報源**といい，また $\mu\in\boldsymbol{P}_T(\Omega)$ のとき $[\Omega,\mu]$ を**定常な複合情報源**という．

次にチャンネルの定常性について論ずる．

(C.3)　$\lambda(Tx,F)=\lambda(x,T^{-1}F)$，$\forall x\in X$，$\forall F\in\mathcal{F}_Y$,

を満たすとき，このチャンネルを**定常**という．$\boldsymbol{C}(X,Y)$ をチャンネル $[X,\lambda,Y]$ の全体，$\boldsymbol{C}_T(X,Y)$ を定常チャンネルの全体とする．また，次の条件

(C.4)　$\lambda(\cdot,[y_i\cdots y_j])\in C(X)$，$\forall[y_i\cdots y_j]$,

を満足するチャンネルを**連続チャンネル**という．

ひとつの整数 $m>0$ が存在して，

(C.5)　任意の $[y_i\cdots y_j]$ $(i\leqq j)$ に対して，$x=\Pi x_k$，$x'=\Pi x_k'\in X$ が $x_k=x_k'(i-m\leqq k\leqq j)$ であるときは，いつでも，

$$\lambda(x,[y_i\cdots y_j])=\lambda(x',[y_i\cdots y_j]),$$

を満足するとき，チャンネルは**（長さ m の）有限記憶**をもつという．チャン

ネルの最終的な条件として

(C.6)　出力源 $(Y, \mathfrak{F}_Y)$ 上に確率測度 ν_0 が存在して，

$$\lambda(x, \cdot) \ll \nu_0, \qquad \forall x \in X$$

を満足するとき $[X, \lambda, Y]$ を**制限** (dominated) チャンネルという．

註　上で定義したチャンネル以外にも，エルゴードチャンネルなどのチャンネルが後で導入される．

以上の各条件の包含関係については，

〈定理 5.4〉　組 X, λ, Y が条件 (C.1) を満たすとき，

$$(\mathrm{C.5}) \Rightarrow (\mathrm{C.4});\quad (\mathrm{C.4}) + (\mathrm{C.6}) \Rightarrow (\mathrm{C.2}).$$

証明　(C.5)⇒(C.4)：これは入力源のメッセージの族 $\{[x_{i-m} \cdots x_j];\ -\infty < i \leq j < \infty\}$ が X の位相基をなすことから容易にいえる．

(C.4)+(C.6)⇒(C.2)：$f \in C(Y)$ に対して，$\bigcup_n \mathfrak{M}_n$ 上の単函数の列 $\{s_n\}$ が存在して，$\sup|s_n(y) - f(y)| \to 0$．このことより，

$$\int s_n(y) \lambda(x, dy) \longrightarrow \int f(y) \lambda(x, dy) \quad (\text{一様収束})$$

条件 (C.4) により $\int s_n(y) \lambda(x, dy) \in C(X)$，したがって $\int f(y) \lambda(x, dy) \in C(X)$．また (C.6) によって，$k(x, y) = \lambda(x, dy)/\nu_0(dy)$ (Radon–Nikodym 微分) が存在して

$$\lambda(x, F) = \int_F k(x, y) \nu_0(dy), \qquad \forall x \in X, \qquad \forall F \in \mathfrak{F}_Y.$$

ところで，任意の $F \in \mathfrak{F}_Y$ に対して $\{f_n\} \subset C(Y)$ が存在して

$$\int_Y |f_n(y) - 1_F(y)| \nu_0(dy) \longrightarrow 0.$$

ゆえに，

$$\begin{aligned}\lambda(x, F) &= \int 1_F(y) k(x, y) \nu_0(dy) \\ &= \lim \int f_n(y) k(x, y) \nu_0(dy)\end{aligned}$$

$$=\lim\int f_n(y)\lambda(x,dy).$$

すなわち，$\lambda(x,F)$ は連続函数の極限であり，したがって，$(X,\mathcal{F}_X)$ 上で可測である．▮

一般に $\mathcal{A}$ を＊代数で単位元 I をもつとし，$\mathcal{B}(\subset\mathcal{A})$ をその＊部分代数(すなわち，部分代数で $b\in\mathcal{B}$ ならば $b^*\in\mathcal{B}$) で，かつ $I\in\mathcal{B}$ とする．$\mathcal{A}$ から $\mathcal{B}$ の上への線形作用素 A が**平均作用素** (averaging operator) であるとは，

$$(Af)^*=Af,\quad A((Af)g)=A(fAg)=AfAg,\quad AI=I,\quad f\geq 0\Rightarrow Af\geq 0.$$

ただし，ある $g\in\mathcal{A}$ (または $g\in\mathcal{B}$) に対して $f=g^*g$ と表されるとき $f\geq 0$ とかく．平均作用素の典型的な例として，ひとつの確率空間 $(\Gamma,\mathcal{F}_\Gamma,P)$ に対する L^∞ を $\mathcal{A}$ とし(これは＊代数)$\mathcal{F}_\Gamma$ の σ 部分集合体 $\mathcal{F}_0$ に関して可測な L^∞ の要素の全体を $\mathcal{B}$ とすると，$\mathcal{B}$ は $\mathcal{A}$ の＊部分代数で I を含み，しかも弱＊位相に関して閉となる．各 $f\in L^\infty$ に対する条件付期待値を $Af=E_p(f|\mathcal{F}_0)$ で表すと，この A が平均作用素である．同様なことが $\mathcal{A}$ が von Neumann 代数，$\mathcal{B}$ が von Neumann 部分代数のときも論ぜられる(第2巻参照)．

有界 Borel 可測函数の全体 $B(X)$, $B(Y)$, $B(\Omega)$ は各点ごとの和，積および複素共役による＊対合によって＊代数であり，さらに"sup-ノルム"を付与することによって可換な C^*-代数となる．また各 $g\in B(X)$ に対して，$g(x,y)=g(x)$ とおくと，g は Ω 上の函数と見なされ，$g\in B(\Omega)$ となり，$B(X)$ は $B(\Omega)$ の C^*-部分代数と見なすことができる．次に，各シフト変換 T を函数に作用させる記号を導入する：$\mathcal{A}=B(X)$, $B(Y)$ または $B(\Omega)$ とし，各 $f\in\mathcal{A}$ に対して $(Tf)(z)=f\circ T(z)$ とおくと，T は $\mathcal{A}$ 上の＊自己同型と見なせ，T は $\mathcal{A}$ から $\mathcal{A}$ 全体への1対1線形写像で，$T(fg)=TfTg$, $Tf^*=(Tf)^*$ が満たされる (5.1 の 9° と同様)．

ここで議論をもとにもどすことにする．各 $\lambda\in\boldsymbol{C}(X,Y)$ に対して，一種の**積分作用素**を

$$(Af)(x)=\int_Y f(x,y)\lambda(x,dy),\qquad x\in X,\ f\in B(\Omega)\qquad (5.12)$$

で定めると，

〈**定理 5.5**〉 作用素 A について次のことが成り立つ.

（1） A は $B(\Omega)\to B(X)$ の平均作用素である.

（2） $\{f_n\}\subset B(\Omega)$, $f\in B(\Omega)$ に対して

$$f_n(\omega)\uparrow f(\omega)(\forall\omega\in\Omega)\Longrightarrow(Af_n)(x)\uparrow(Af)(x)(\forall x\in X).$$

（3） $AT=TA\Longleftrightarrow\lambda\in\boldsymbol{C}_T(X,Y)$.

ここで,

$$(Tf)(x,y)=f(Tx,Ty),\quad (Tg)(x)=g(Tx)\quad(\forall f\in B(\Omega),\forall g\in B(X))$$

とする.

証明 （1）: $A((Af)g)=(Af)(Ag)$ $(\forall f,g\in B(\Omega))$ を示す（他は容易である).

$$A((Af)g)(x)=\int(Af)(x)g(x,y)\lambda(x,dy)$$
$$=(Af)(x)\int g(x,y)\lambda(x,dy)=(Af)(x)(Ag)(x).$$

（2）: これは単調収束定理から明らかである.

（3）: ($\Leftarrow$) は各 $f\in B(\Omega)$ に対して

$$(ATf)(x)=\int(Tf)(x,y)\lambda(x,dy)=\int f(Tx,Ty)\lambda(x,dy)$$
$$=\int f(Tx,y)\lambda(x,T^{-1}dy)=\int f(Tx,y)\lambda(Tx,dy)$$
$$=(Af)(Tx)=(TAf)(x)\quad(\forall x\in X)$$

からいえ, 逆の ($\Rightarrow$) は, $F\in\mathcal{F}_Y$ に対して $f(x,y)=1_F(y)$ として, 上と同様の積分の式を計算すればすぐに得られる. ▮

この定理の（1),（2）を満たす線形作用素 A の全体を $\boldsymbol{A}(X,Y)$ で表し,（3）を満たす A の全体を $\boldsymbol{A}_T(X,Y)$ で表す. この記号のもとで次の定理がいえる.

〈**定理 5.6**〉 $A\in\boldsymbol{A}(X,Y)$ に対して $\lambda\in\boldsymbol{C}(X,Y)$ が一意に存在して (5.12) を満たし, 対応 $A\longleftrightarrow\lambda$ は $\boldsymbol{A}(X,Y)$ と $\boldsymbol{C}(X,Y)$ の間に1対1の対応を与え

る．また，この対応は $\boldsymbol{A}_T(X, Y)$ と $\boldsymbol{C}_T(X, Y)$ の間の1対1対応でもある．

これの証明は容易である．以下の節では，λ に対応する平均作用素を A_λ で表す．

$\boldsymbol{A}(X, Y)$ は $B(\Omega)\to B(X)$ である有界線形作用素のなす線形空間の凸部分集合である．$\boldsymbol{A}(X, Y)$ の端点の全体を $\mathrm{ex}\,\boldsymbol{A}(X, Y)$ で表す．

この節の議論は主に Umegaki [U.6] によっている．この節と次の節の議論の一部は量子系にも拡張されている [O.3, 4] が，それらに関しては第2巻で詳しく説明する．

5.4 チャンネルのエルゴード特性

この節ではチャンネルの族 $\boldsymbol{C}(X, Y)$，$\boldsymbol{C}_T(X, Y)$ について考える．ここで 5.3 節と同様，$X=A^Z$, $Y=B^Z$, $\Omega=X\times Y=(A\times B)^Z$ である．まず，いくつかの記号を導入しよう．

各 $\mu\in\boldsymbol{P}(X)$ に対して

$$\mu'(E)=\int_X \lambda(x,\ E)d\mu(x),\qquad E\in\mathfrak{F}_Y, \tag{5.13}$$

$$\mu''(F)=\int_X \lambda(x, F_x)d\mu(x),\qquad F\in\mathfrak{F}_\Omega. \tag{5.14}$$

ここで $F_x=\{y\in Y;\ (x, y)\in F\}$，これは $F_x\in\mathfrak{F}_Y$．

次の命題が成立する．

1° $\mu\in\boldsymbol{P}_T(X)$，$\lambda\in\boldsymbol{C}_T(X, Y)\Longrightarrow\mu'\in\boldsymbol{P}_T(Y)$，$\mu''\in\boldsymbol{P}_T(\Omega)$．

実際，$F=C\times D$ $(C\in\mathfrak{F}_X, D\in\mathfrak{F}_Y)$ に対して

$$\begin{aligned}\mu''(T^{-1}(C\times D))&=\mu''(T^{-1}C\times T^{-1}D)\\&=\int_{T^{-1}C}\lambda(x, T^{-1}D)d\mu(x)=\int_C\lambda(Tx, D)d\mu(Tx)\\&=\int_C\lambda(y, D)d\mu(y)=\mu''(C\times D)\end{aligned}$$

$(=\mu''(C, D)$ とかくこともある$)$．

一般の $F\in\mathcal{F}_\Omega$ に対しては，$\mathcal{F}_\Omega=\sigma\{C\times D;\ C\in\mathcal{F}_X,\ D\in\mathcal{F}_Y\}$ からいえる．$\mu'\in\boldsymbol{P}_T(Y)$ は μ'' の特別な場合である．

定常チャンネル $\lambda\in\boldsymbol{C}_T(X,Y)$ が**エルゴード的**であるとは

(C.7) $\mu\in\boldsymbol{P}_E(X)\Longrightarrow\mu''\in\boldsymbol{P}_E(\Omega)$,

であるときをいい，このとき，$[X,\lambda,Y]$ を**エルゴード情報源**という．明らかに，$\mu''\in\boldsymbol{P}_E(\Omega)$ ならば $\mu'\in\boldsymbol{P}_E(Y)$ である．エルゴード的な λ の全体を $\boldsymbol{C}_E(X,Y)$，あるいは単に $\boldsymbol{C}_E$ で表す．なお，$\lambda\in\boldsymbol{C}_E(X,Y)$，あるいは $[X,\lambda,Y]$ を**エルゴード・チャンネル**という．

このことも平均作用素 A (5.3 節参照) を用いて論ずると次のように表される．A は $B(\Omega)\to B(X)$ の線形作用素であり，A の共役作用素 A^* は $B(X)^*\to B(\Omega)^*$ となる線形作用素である．A^* の定義域を $\boldsymbol{P}(X)$ および $\boldsymbol{P}_T(X)$ に制限すると，$\mu\in\boldsymbol{P}(X)$ に対して，$\mu''=A^*\mu$ であり，A が $TA=AT$ (すなわち，A に対応する λ が $\lambda\in\boldsymbol{C}_T(X,Y)$) で $\mu\in\boldsymbol{P}_T(X)$ ならば $A^*\mu=\mu''\in\boldsymbol{P}_T(\Omega)$.

エルゴード・チャンネルの特性化定理を示すために，'mod T' および 'm-ex' の定義を与える．$\lambda,\lambda'\in\boldsymbol{C}_T(X,Y)$ に対して $\lambda=\lambda'\ (\mathrm{mod}\ T)$ であるとは，任意の $\mu\in\boldsymbol{P}_T(X)$ と任意の $F\in\mathcal{F}_Y$ に対して $\lambda(x,F)=\lambda'(x,F)\ \mu$-a.e. $x\in X$ のことである．このことは，任意の $\mu\in\boldsymbol{P}_T(X)$，任意の $f\in B(\Omega)$ に対して，$(A_\lambda f)(x)=(A_{\lambda'}f)(x)\ \mu$-a.e. であることと同等である．この記号を用いて，$\lambda\in\boldsymbol{C}_T(X,Y)$ が $\boldsymbol{C}_T(X,Y)$ における**端点**であるとは，$\lambda=\alpha\lambda_1+(1-\alpha)\lambda_2\ (\mathrm{mod}\ T)$ $(\lambda_1,\lambda_2\in\boldsymbol{C}_T(X,Y),\ 0\leqq\alpha\leqq1)$ ならばつねに $\lambda=\lambda_1=\lambda_2$ であることをいう．この意味での $\boldsymbol{C}_T$ の端点の全体を m-ex $\boldsymbol{C}_T(X,Y)$ で表す．T-定常な平均作用素の全体 $\boldsymbol{A}_T$ においても，同様に m-ex $\boldsymbol{A}_T$ が定義される．

〈定理 5.7〉 定常チャンネル $[X,\lambda,Y]$ において，次の条件は同等である：

(1) $[X,\lambda,Y]$ はエルゴード・チャンネルである．

(2) $\lambda\in m\text{-ex}\,\boldsymbol{C}_T(X,Y)$.

(3) $A_\lambda\in m\text{-ex}\,\boldsymbol{A}_T(X,Y)$.

（4）　任意の対 $f, g \in C(\Omega)$ に対して

$$\lim_{n\to\infty} A(f_n g)(x) = \lim_{n\to\infty}(Af_n)(x)(Ag)(x), \quad \mu\text{-a.e. } x \in X, \ \forall \mu \in \boldsymbol{P}_E(X).$$

この定理は，情報理論における有用性もさることながら，エルゴード性の特性定理の一般化として，より意味がある．この定理は Nakamura [N. 10] と Umegaki [U. 6] が同じ頃（1968 年），独立に，多少の差異はあるが証明したものである．

〈定理 5.7〉の条件（4）は次の形にかき直すことができる：

（5）　任意のメッセージ $E, F \subset Y$ に対して

$$\lim \frac{1}{n}\sum_{k=0}^{n-1}(\lambda(x, E\cap T^{-k}F) - \lambda(x, E)\lambda(x, T^{-k}F)) = 0,$$

$$\mu\text{-a.e. } x \in X, \ \forall \mu \in \boldsymbol{P}_E(X).$$

また，この極限の式においてさらに E, F を $E, F \in \mathcal{F}_Y$ と一般な可測集合に置き換えることも可能である．この条件は Adler [A. 4] によって導入された“漸近的収束”条件：

$$\lim_{n\to\infty}(\lambda(x, E\cap T^{-n}F) - \lambda(x, E)\lambda(x, T^{-n}F)) = 0, \quad x \in X$$

を含んでいる．

註　チャンネル $[X, \lambda, Y]$ がいつエルゴード性をもつかという議論は Khinchin の論文以来，種々に論ぜられてきた．Khinchin は有限記憶の条件（C. 5）のもとでこれを論じたが，証明にギャップがあることを K. Takano [T. 1] が指摘し，同氏はチャンネルが有限歩（m 歩）従属な条件：

（C. 8）　自然数 $m>0$ が存在して，任意の整数の組 n, r, s, t $(n \leqq r \leqq s \leqq t)$ に対して $s-r>m$ ならば，出力上のメッセージ $E_{n,r}=[y_n \cdots y_r]$ と $E_{s,t}=[y_s \cdots y_t]$ に対して，つねに

$$\lambda(x, E_{n,r}\cap E_{s,t}) = \lambda(x, E_{n,r})\lambda(x, E_{s,t}), \quad \forall x \in X,$$

を満足するとき，定常な $[X, \lambda, Y]$ がエルゴード性をもつことを示した．Takano の仕事をきっかけとして，以後約10年間で，上述のように，これらを含む数学的に明解な条件が見出された．

ここで，有限記憶 m をもち，m 歩従属であるチャンネル $[X, \lambda, Y]$ を構成しておこう．A, B をある有限集合とし，$X=A^Z$，$Y=B^Z$ とおく．任意の

$x\in A\times\cdots\times A\equiv A^m$ に対して，B^m 上の確率分布 $p(x|\cdot)$ が与えられたとする．各 $x\in X$ と各 $[y_s\cdots y_t]$ $(s\leqq t)$ に対し，

$$\lambda(x,[y_s\cdots y_t])=\prod_{k=0}^{t-s}p((x_{s+k-m},\cdots,x_{s+k-1})|(y_{s+k-m},\cdots,y_{s+k-1}))$$

で与えられる測度 $\lambda(x,\cdot)$ を $\mathcal{F}_Y$ 上へ一意に拡張し，それも同じ $\lambda(x,\cdot)$ で表す．このとき，$[X,\lambda,Y]$ は定常で，有限記憶 m をもち，m 歩従属なチャンネルである．以下これを証明しよう．

定義より，

$$\begin{aligned}
&\lambda(x,[y_r\cdots y_s]\cap[y_{s+m+1}\cdots y_t])\\
&=\sum\nolimits_{[y_{s+1}\cdots y_{s+m}]\in\mathcal{M}_m(Y)}\lambda(x,[y_r\cdots y_s y_{s+1}\cdots y_{s+m}y_{s+m+1}\cdots y_t])\\
&=\sum\prod_{j=0}^{s-r}p((x_{r+j-m},\cdots,x_{r+j})|(y_{r+j-m},\cdots,y_{r+j}))\\
&\quad\cdot p((x_{s+1},\cdots,x_{m+s})|(y_{s+1},\cdots,y_{s+m})\\
&\quad\cdot\prod_{k=1}^{t-s-2m+1}p((x_{s+m+k},\cdots,x_{s+2m+k-1})|(y_{s+m+k},\cdots,y_{s+2m+k-1}))\\
&=\prod_{j=0}^{s-r}p((x_{r+j-m},\cdots,x_{r+j})|(y_{r+j-m},\cdots,y_{r+j}))\\
&\quad\cdot\sum p((x_{s+1},\cdots,x_{s+m})|(y_{s+1},\cdots,y_{s+m}))\\
&\quad\cdot\prod_{k=0}^{t-s-2m+1}p((x_{s+m+k},\cdots,x_{s+2m+k-1})|(y_{s+m+k},\cdots,y_{s+m+k-1}))\\
&\equiv\alpha\cdot\beta\cdot\gamma\quad(\text{とおく}).
\end{aligned}$$

各 α,β,γ は

$$\alpha=\lambda(x,[y_r\cdots y_s]),\qquad\beta=1,\qquad\gamma=\lambda(x,[y_{s+m+1}\cdots y_t])$$

となる．よって λ は m 歩従属である．

5.5 相互エントロピーと伝送容量

入力情報源 $[X,\mu]$，$\mu\in\boldsymbol{P}_T(X)$ のエントロピー $S(\mu)$ は，その定義から，一見非線形なスカラー量に思えるが，実はこれが線形性（アフィン性）をもっていることが示せる：

〈定理 5.8〉　任意の $\mu, \nu \in \boldsymbol{P}_T(X)$ と実数 $\alpha, \beta \geqq 0$, $\alpha+\beta=1$ に対して

$$S(\alpha\mu+\beta\nu)=\alpha S(\mu)+\beta S(\nu).$$

証明　$\alpha, \beta>0$ としてよい．長さ n のメッセージのエントロピー S_n について，

$$\begin{aligned}S_n(\alpha\mu+\beta\nu)&=-\sum(\alpha\mu(C)+\beta\nu(C))\log(\alpha\mu(C)+\beta\nu(C))\\&=-\alpha\sum\mu(C)\log\mu(C)-\beta\sum\nu(C)\log\nu(C)\\&\quad-\alpha\sum\mu(C)\log\{\alpha+\beta\nu(C)/\mu(C)\}\\&\quad-\beta\sum\nu(C)\log\{\beta+\alpha\mu(C)/\nu(C)\}. \qquad (5.15)\end{aligned}$$

ここで，$\sum$ は $C\in\mathcal{M}_n$ にわたるが，$\mu(C)=0$ および $\nu(C)=0$ である項は除外して考えることにする．さらに，

$$\log\alpha\leqq\log(\alpha+\beta\nu(C)/\mu(C))\leqq\log\alpha+\beta\nu(C)/\alpha\mu(C),$$

$$\alpha\log\alpha\leqq\sum\alpha\mu(C)\log(\alpha+\beta\nu(C)/\mu(C))\leqq\alpha\log\alpha+\beta$$

より，(5.15) の両辺に $1/n$ を掛け，$n\to\infty$ をとると，

$$\begin{aligned}S(\alpha\mu+\beta\nu)&\equiv\lim_{n\to\infty}\frac{1}{n}S_n(\alpha\mu+\beta\nu)\\&=\alpha\lim_{n\to\infty}\frac{1}{n}S_n(\mu)+\beta\lim_{n\to\infty}\frac{1}{n}S_n(\nu)\\&=\alpha S(\mu)+\beta S(\nu). \quad \blacksquare\end{aligned}$$

ここで，伝送速度という概念を導入するために，新たなエントロピーが必要となる．一般に，2つの事象系

$$\tilde{C}\equiv(\{c_i\}_{i=1}^n, \{p(c_i)\}_{i=1}^n), \qquad \tilde{D}\equiv(\{d_j\}_{j=1}^m, \{p(d_j)\}_{j=1}^m)$$

が与えられ，それらの複合事象系 $\tilde{C}\times\tilde{D}\equiv(\{(c_i, d_j)\}, \{p(c_i, d_j)\})$ が与えられたとき，

$$\begin{aligned}I_p(\tilde{C}, \tilde{D})&\equiv\sum_{i,j}p(c_i, d_j)\log\frac{p(c_i|d_j)}{p(c_i)}\\&=\sum_{i,j}p(c_i, d_j)\log\frac{p(c_i, d_j)}{p(c_i)p(d_j)}\\&=-\sum_i p(c_i)\log p(c_i)+\sum_{i,j}p(c_i, d_j)\log\frac{p(c_i, d_j)}{p(d_j)}\end{aligned}$$

$$= -\sum_j p(d_j)\log p(d_j) + \sum_{i,j} p(c_i, d_j)\log \frac{p(c_i, d_j)}{p(c_i)}.$$

したがって，$I_p(\tilde{C}, \tilde{D}) = I_p(\tilde{D}, \tilde{C})$ であり，これを $\tilde{C}$ と $\tilde{D}$ **の間の相互エントロピー**（または，**相互情報量**）という．

$\log(p(c_i|d_j)/p(c_i))$ は d_j を知ったという条件のもとで，$c_i \in \tilde{C}$ の情報量を表しており，$I_p(\tilde{C}, \tilde{D})$ はそれを $\tilde{C} \times \tilde{D}$ にわたって分布 p で平均したものである．また，$p(c_i, d_j) = p(c_i)p(d_j)$ $(\forall i, j)$ のとき，$\tilde{C}$ と $\tilde{D}$ は p に関して独立であるといわれるが，このとき $I_p(\tilde{C}, \tilde{D}) = 0$ となり，よって，相互エントロピーは $\tilde{C}$ と $\tilde{D}$ の独立性の程度を表す統計量と考えることもできる．

さて，ここで再び定常チャネル $[X, \lambda, Y]$ の話にもどる．定常な入力情報源の測度 μ が設置されたとき，複合情報源 $[\Omega, \mu'']$ がただちに構成され，入力と出力の各情報源はともに定常で，それらは $[\Omega, \mu'']$ の部分事象と見なされる．したがって，上述の相互エントロピーがこの場合も議論できる．$X = A^Z$ および $Y = B^Z$ 上の長さ n のメッセージ全体 $\mathcal{M}_n(X)$，$\mathcal{M}_n(Y)$ において，任意の $C \in \mathcal{M}_n(X)$ に対して，

$$S_n(\mu; Y|C) = -\sum_{D \in \mathcal{M}_n(Y)} \mu''(C \times D)\log \frac{\mu''(C \times D)}{\mu(C)}$$

とし，これのすべての $C \in \mathcal{M}_n(X)$ についての和

$$S_n(\mu; Y|X) = \sum_{C \in \mathcal{M}_n(X)} S_n(\mu; Y|C)$$

は，長さ n のメッセージのいずれかを送ったという条件のもとで，長さ n のメッセージを得るための条件付エントロピーである．入力および出力情報源における長さ n のメッセージ間の相互エントロピーは次の式で与えられる．

$$I_{\mu''}(\mathcal{M}_n(X),\ \mathcal{M}_n(Y)) = S_n(\mu') - S_n(\mu; Y|X).$$

ただし，$S_n(\mu') = -\sum_{D \in \mathcal{M}_n(Y)} \mu'(D)\log \mu'(D)$ である．

さらに，簡単な計算より，

$$I_{\mu''}(\mathcal{M}_n(X),\ \mathcal{M}_n(Y)) = S_n(\mu) + S_n(\mu') - S_n(\mu'')$$

となり，したがって，$I_{\mu''} \geqq 0$ である．さらに，

$$I(\mu; \lambda) \equiv \lim_{n\to\infty} \frac{1}{n} I_{\mu''}(\mathcal{M}_n(X), \mathcal{M}_n(Y))$$

$$= S(\mu) + S(\mu') - S(\mu'') \tag{5.16}$$

が存在する．これは，入力情報源 $[X, \mu]$ のメッセージがチャンネル λ を通して伝送されたときに得られる1文字当たりの平均情報量であり，通信の伝送機構における伝送の正確さの度合いを表す比率でもある．これを**伝送速度**，あるいは入力・出力間の**相互エントロピー**（**相互情報量**）とよぶ．

等式（5.16）と〈定理 5.8〉より $I(\mu;\lambda)$ の線形性が示せる：

〈定理 5.9〉　任意の $\mu, \nu \in \boldsymbol{P}_T(X)$ と実数 $\alpha, \beta > 0$，$\alpha + \beta = 1$ に対して，

$$I(\alpha\mu + \beta\nu;\lambda) = \alpha I(\mu;\lambda) + \beta I(\nu;\lambda).$$

以上，伝送速度を論じてきたが，その際，チャンネル $[X, \lambda, Y]$ の定常性以外は何も制約を設けなかった．数学的には，ここまでの議論は，2つの情報源の間の一種の測度の変換に関する一般論を展開したことと同じである．しかしながら，実際の通信においては，有限個のアルファベットの上で行われるため，通信機構の統計的構造として何らかの有限性をもたせるという制約をおくことによって，議論をより精密なものにしうるのである．

$\boldsymbol{P}_T(X)$ は $C(X)$ の共役空間 $C(X)^*$ の単位球 $C(X)_1{}^*$ の弱*位相に関してコンパクトな凸部分集合であることを使うと，函数 $I(\cdot;\lambda)$；$\boldsymbol{P}_T(X) \to [0, \infty)$ の連続性がわかる：

〈定理 5.10〉　チャンネル $[X, \lambda, Y]$ が定常で有限記憶をもち，有限歩従属であれば，$I(\mu;\lambda)$ は $\boldsymbol{P}_T(X)$ 上で上半連続である．

この定理は Breiman [B. 9] が証明したものであり，その証明は数学的にかなり興味あるものであるが，ここでは割愛する．さて，

$$C_s \equiv \sup\{I(\mu;\lambda);\ \mu \in \boldsymbol{P}_T(X)\},$$

$$C_e \equiv \sup\{I(\mu;\lambda);\ \mu \in \boldsymbol{P}_E(X)\ \text{かつ}\ \mu'' \in \boldsymbol{P}_E(\Omega)\}$$

を，それぞれ定常チャンネル λ に関する，**（定常）伝送容量**，**エルゴード伝送容**

量という．これらは，チャンネルを通して通信文をできるだけ正確に送るための指標となる数値である．C_e は，以下の定理で示すように，チャンネル λ がエルゴード的 (i.e.，$\lambda\in \boldsymbol{C}_E(X, Y)$) のときとくに意義がある．

〈**定理 5.11**〉 定常チャンネル $[X, \lambda, Y]$ に関して，C_e, C_s が存在して，

$$0\leqq C_e\leqq C_s\leqq \log ab.$$

さらに，$[X, \lambda, Y]$ が有限記憶をもち，有限歩従属であれば

$$C_e=C_s$$

となる．ここで，a, b は各々 A, B の要素の個数である．

証明 $C_s\geqq C_e\geqq 0$ は定義より明らか．$\Omega=X\times Y=(A\times B)^{\boldsymbol{Z}}$ であるから，$A\times B$ の要素の個数は ab で，$\mu\in \boldsymbol{P}_T(X)$ に対して，各 $[X, \mu]$, $[Y, \mu']$, $[\Omega, \mu'']$ のエントロピー $S(\mu), S(\mu'), S(\mu'')$ は $0\leqq S(\mu)\leqq \log a$, $0\leqq S(\mu')\leqq \log b$, $0\leqq S(\mu'')\leqq \log ab$ を満たす．したがって，

$$0\leqq I(\mu;\lambda)=S(\mu)+S(\mu')-S(\mu'')\leqq \log ab$$

となる．よって，定義より，

$$0\leqq C_e\leqq C_s\leqq \log ab<+\infty.$$

次に，$\lambda\in \boldsymbol{C}_E(X, Y)$ のとき，$C_e=C_s$ を示す．以下の計算において，$I(\mu;\lambda)$ の $\boldsymbol{P}_T(X)$ 上の線形性（〈定理 5.9〉）と上半連続性（〈定理 5.10〉），および関係式 $\boldsymbol{P}_E(X)=\mathrm{ex}\,\boldsymbol{P}_T(X)$ を用いる．

$$\begin{aligned}
C_e\leqq C_s &\equiv \sup\{I(\mu;\lambda);\ \mu\in \boldsymbol{P}_T(X)\}\\
&=\sup\{I(\textstyle\sum \alpha_k\mu_k;\lambda);\ \mu_k\in \boldsymbol{P}_E(X);\ \sum \alpha_k=1, \alpha_k\geqq 0\}\\
&=\sup\{\textstyle\sum \alpha_k I(\mu_k;\lambda);\ \mu_k\in \boldsymbol{P}_E(X),\ \sum \alpha_k=1, \alpha_k\geqq 0\}\\
&\leqq \sup\{I(\mu;\lambda);\ \mu\in \boldsymbol{P}_E(X)\}\\
&\equiv C_e.
\end{aligned}$$

ゆえに，$C_e=C_s$. ▮

さらに，次の定理が成立する．

〈**定理 5.12**〉 チャンネル $[X, \lambda, Y]$ が定常で，有限記憶をもち，有限歩従

属であれば，エルゴード情報源 $[X,\mu]$ (すなわち，$\mu\in\boldsymbol{P}_E(X)$) が存在して，

$$C_e=I(\mu;\lambda).$$

証明　〈定理 5.10〉と〈定理 5.11〉によって，$\mu\in\boldsymbol{P}_T(X)$ が存在して，$C_e=I(\mu;\lambda)$ となる．今

$$\boldsymbol{P}_0=\{\mu\in\boldsymbol{P}_T(X);\ I(\mu;\lambda)=C_e\}\quad(\neq\emptyset)$$

とおくと，$\boldsymbol{P}_0$ は，$I(\cdot;\lambda)$ の上半連続性より $\boldsymbol{P}_T(X)$ の凸な閉部分集合である．ゆえに，$\boldsymbol{P}_0$ は弱＊コンパクトになり，Krein–Milman の定理（付章5節）が使えて，$\mu\in\mathrm{ex}\,\boldsymbol{P}_0$ が存在する．これが $\mu\in\mathrm{ex}\,\boldsymbol{P}_T(X)$ であることを示せばよい．もし，ある $p,q\in\boldsymbol{P}_T(X)$，$\alpha\in(0,1)$ に対して $\mu=\alpha p+(1-\alpha)q$ とかければ，

$$\begin{aligned}C_e=I(\mu;\lambda)&=I(\alpha p+(1-\alpha)q)\\&=\alpha I(p;\lambda)+(1-\alpha)I(q;\lambda)\end{aligned}$$

となるから，$I(p;\lambda)=I(q;\lambda)=C_e$，すなわち，$p,q\in\boldsymbol{P}_0$ となる．ゆえに $\mu=p=q$．▮

この定理は，伝送容量 $C_e(=C_s)$ にちょうど一致するエルゴード情報源 μ の存在が実証されたので，情報の伝送において高い利用価値がある．さらに，この μ がどのような（具体的な）分布であるかを調べることは興味あることである．

"$C_e=C_s$" の証明は Khinchin [K. 10] 以後，多数の数学者によって論じられたが（本書では Breiman [B. 9] の方法を採用），Parthasarathy [P. 2] は有限記憶の仮定なしに証明した．彼は，定常情報源の確率測度 μ が Kryloff–Bogolioubov 流の積分表示によるエルゴールド分解を行い，非常に明解な証明を与えた．これは数学的に興味ある証明であるが，エルゴード分解に関する準備が必要なので，ここでは割愛する．

なお，Parthasarathy の証明によると，チャンネルのエルゴード性の仮定のみで有限性の制約は必要としないことがわかる．

この節の最後に，一般のコンパクト Hausdorff 空間 X,Y において，$\mu\in\boldsymbol{P}(X)$ がチャンネル λ を通して $\mu'\in\boldsymbol{P}(Y)$ に変化したとすると，この

μ' と，μ と μ' の合成測度 μ'' は (5.13) と (5.14) によって与えられるが，この場合の相互エントロピーを与えておこう．今，μ と μ' の直積測度を $\mu_0''=\mu\otimes\mu'$ とおくと，$\mu''\ll\mu_0''$ となるとき，R–N 微分 $d\mu''/d\mu_0''$ が存在する．この場合，μ と μ' の相互エントロピーは

$$I(\mu:\lambda)=\int_\Omega\left(\frac{d\mu''}{d\mu_0''}\log\frac{d\mu''}{d\mu_0''}\right)d\mu_0''\quad\left(\equiv S(\mu''|\mu_0'')\right)\qquad(5.17)$$

で与えられる．この表現は非常に有用である．

5.6 符号化の定理

一般に，通常の伝送機構では，入力情報源から信号（メッセージ，文字列）を送り続けるとき，受信側としてある時刻 t でひとつの文字を受け取る確率は，t までに送り込まれた文字列のみに依存し，t 以後にどんな信号が送られたかにはよらない．つまり，受信者として t 以後の信号を予測しないことになる．これをチャンネルを使って表現すると，次のようになる．

チャンネル $[X,\lambda,Y]$ が**非予測**であるとは，文字 $b\in B$ に対して，メッセージ $[y_t]$ が $y_t=b$ であるとき，$[y_t=b]$ とかくことにすると，

(C.8) 文字列の対 $x=\Pi\,x_k$, $x'=\Pi\,x_k'\in X$ が $x_k=x_k'$ $(\forall k\leqq t)$ であれば，つねに

$$\lambda(x,[y_t=b])=\lambda(x',[y_t=b]),$$

が成立する場合をいう．これまでのチャンネルの議論ではこの条件は全く不要であったが，以下の議論ではこの条件を仮定したうえで行われる．**この節では，チャンネル $[X,\lambda,Y]$ は定常で有限記憶 m をもち，エルゴード性および非予測性をももつと仮定する**（〈仮定 C〉）．

まず，次の有用な定理から述べよう．

〈定理 5.13〉 Feinstein の定理：〈仮定 C〉のもとで任意の $\varepsilon>0$, $\alpha\in(0,C_e)$ に対して，自然数 $n=n(\varepsilon,\alpha)$ と N を適当に選ぶと，つねに長さが $m+n$ の N 個のメッセージ $U_1,\cdots,U_N(\in\mathcal{M}_{m+n}(X))$ と N 個の $V_1,\cdots,V_N\subset Y$ が存在して，次の条件を満たす．

（1）　$V_i \cap V_j = \emptyset \quad (i \neq j)$.

（2）　$x \in U_j \Longrightarrow \lambda(x, V_j) > 1-\varepsilon \quad (\forall j)$.

（3）　$N > e^{n(C_e - \alpha)}$.

この定理の情報論的意味は次のようである：各 j に対して，$\lambda(x, V_j) > 1-\varepsilon$ $(\forall x \in U_j)$ であり，ε は十分小さいので，メッセージ U_j が入力情報源から送られたとき，出力側では1に近い確率で V_j に含まれる Y のメッセージが得られ，しかもそれが他の V_k に含まれているものとは明確に区別できることがわかり，また，逆に V_j が受信されたことを知れば，1に近い確率で送信されたメッセージ U_j を知ることができる．

なお，この定理の証明は文献 [K.19] を参照されるとよい．

さて，入力情報源 $X = A^Z$ の文字の列に対し，新たな記号の列を1対1に対応させることを**符号化** (coding) という．これまでは，チャンネル $[X, \lambda, Y]$ の入力に接合されたアルファベットは情報源の送信のアルファベットと同じものとして議論を進めてきた．しかし，実際には，多くの場合この2つのアルファベットは異なっている．たとえば，日本語で電報を打つ場合を考えると，まず発信人が頼信紙にカタカナで文章（メッセージ）をかき，それを電報局が電信記号に符号化して相手の電報（受信）局に送る．この場合，はじめに発信人が用いたアルファベットと電信記号のアルファベットは一般に異なる．

そこで，A_0 を A と異なるアルファベットとし，メッセージ空間 $X_0 = A_0^Z$ を考えよう．定常情報源 $[X_0, \mu]$ から出されるメッセージが定常チャンネル λ を通して送られる場合を考える．このとき，X_0 のメッセージを X のそれに翻訳（符号化）する必要がある．X_0 の要素を X の要素に1対1に対応させる規則を φ で表すと，符号化とはこの1対1の写像 φ; $X_0 \longrightarrow X$ のことである．したがって，X_0 を入力，X を出力のメッセージ空間と考えることができ，

$$\lambda_\varphi(\xi, U) = 1_U(\varphi(\xi)), \quad \forall \xi \in X_0, \forall U \in \mathscr{F}_X$$

とおくと，$[X_0, \lambda_\varphi, X]$ は（雑音のない）ひとつのチャンネルとなる．$\lambda_\varphi(\xi, U) = 1$ のとき，$\xi \in X_0$ が $\varphi(\xi)$ に符号化されたことを知る．

このように，符号化とは $X_0=A_0^{\mathbf{Z}} \longrightarrow X=A^{\mathbf{Z}}$ である1対1写像であるが，数学的にはそのような写像はいくらでも構成できる．

ところで，X_0 と X の要素はともにアルファベットの無限列であるが，現実に通信などに現れるのは有限の長さをもったメッセージであるので，有限の文字列に限定した定式化を行う方がより細かな確率計算をするうえで便利である．各 $x_k \in A$ をひとつの符号と考え，これに A_0 の文字の有限列 $\{\xi_j\} \subset A_0$ が対応している．すなわち，A_0 の文字を使ったひとつの特定な文章を A の1文字に対応するという関係をまず与える．次に，これを拡大し，A_0 の一定の長さ以下に押さえられた文章の集団（つねに有限個）の各々に A の文字のやはり一定の長さに押さえられた文字の集団を対応させるという対応関係をつくる．いいかえると，A_0 の文字の長さ m 以下の文章の集団を A の文字の長さが n 以下の文章の集団に対応させる規則をつくるのである．これが現実的な符号化である．

〈定理 5.14〉 Shannon の第一定理：〈仮定 C〉のもとで，エルゴールド情報源 $[X_0, \mu_0]$ のエントロピー $S(\mu_0)$ が $S(\mu_0)<C_e$ (C_e はチャンネル $[X, \lambda, Y]$ のエルゴールド伝送容量）を満たすとき，$\varepsilon>0$ に対して十分大きな n をとれば，$[X_0, \mu_0]$ から出される長さ n のメッセージ $\alpha(\in \mathcal{M}_n(X))$ を長さ $n+m$ (m は λ の記憶の長さ）のメッセージ $U(\in \mathcal{M}_{n+m}(X))$ に符号化し，この U をチャンネル $[X, \lambda, Y]$ を通して送信したとき，出力側 Y で得られるメッセージ V によって，送られたメッセージ α を $1-\varepsilon$ より大きな確率で知ることができる．

符号化を与える写像を $\varphi: X_0 \to X$ とし，

$$\lambda_\varphi(\xi, V)=\lambda(\varphi(\xi), V), \quad \forall \xi \in X_0, \ \forall V \in \mathcal{F}_Y$$

とすると，新たなチャンネル $[X_0, \lambda_\varphi, Y]$ が得られ，これを使うと，上の定理の結論は $\lambda_\varphi(\xi, V)>1-\varepsilon\ (\forall \xi \in \alpha)$ という式で表される．つまり，$S(\mu_0)<C_e$ ならば，受信されたメッセージから送信されたメッセージをいくらでも高い精

度で推定することができるのである.

〈定理 5.15〉 Shannon の第二定理:〈定理 5.14〉の条件のもとで，伝送速度が $S(\mu_0)$ にいくらでも近くなるような符号化を選ぶことができる.

$S(\mu_0)<C_e$ なる伝送容量をもったチャンネルでは，$[X_0, \mu_0]$ のメッセージを誤りなく符号化する方法が存在することを〈定理 5.15〉は主張しているのである.

これらの定理の証明は文献 [B.7, F.1,2, K.12,14] にある.

第6章　信　号　解　析

信号解析は，元来通信工学の分野であるが，情報理論が数学分野として発展するにつれ，厳密な数学的方法が導入されて，Fourier 解析や函数解析の延長上の理論として重要性が増している．この章の 6.1 節では，信号解析の基本的手段となる Fourier 変換に関する最も必要とされる基本的な諸定理について説明する．6.2 節では，信号概念について述べ，種々な信号函数の数式表示を行い，信号函数の解析にとって不可欠な標本化定理に及ぶ．6.3 節では，ランダムな信号の解析を行うため函数解析の領域で著名な Mercer の定理より始め，Karhunen-Loéve 展開について論ずる．最終の 6.4 節では，Heisenberg 正準交換関係を von Neumann 流に作用素論的に説明し，これと同等の関係が信号のスペクトルのゆらぎと観測時間のゆらぎとの間にも成立することを示す．また，この節で，不確定性関係とエントロピーの係わりについても調べる．

6.1　Fourier 変換

この節では，実数直線 $\boldsymbol{R}=(-\infty,\infty)$ 上の Fourier 変換に関する最も基本的な諸定理について述べる．以下，$\boldsymbol{R}$ 上の函数の生成する Banach 空間を $L^p\equiv L^p(\boldsymbol{R})$ $(p\geq 1)$ で，およびコンパクト台をもつ連続函数のなす線形空間を $C_{00}\equiv C_{00}(\boldsymbol{R})$ で表す．

6.1.1　たたみ込み $f*g$

$\boldsymbol{R}$ 上の可測函数 f, g に対して

$$f*g(t)=\int_{\boldsymbol{R}} f(t-s)g(s)d(s)$$

が存在するとき，$f*g$ を f, g の**たたみ込み**という．

以下 L^p 空間のノルムを $\|\cdot\|_p$ で表す．

〈定理 6.1〉　次の（1），（2），（3）の各々の f,g に対してつねに $f*g$ が存在し，$f*g=g*f$ かつ次の各関係式が成立する．

（1）　$f,g\in L^1 \Longrightarrow \|f*g\|_1\leq\|f\|_1\|g\|_1$.

（2）　$f\in L^1$，$g\in L^2 \Longrightarrow \|f*g\|_2\leq\|f\|_1\|g\|_2$.

（3）　$f\in L^p$ $(p\geq 1)$，$g\in L^q$ $(1/p+1/q=1) \Longrightarrow f*g\in C_0$.

〈補題 6.2〉　$f\in L^p$ に対して

$$\int_R |f(t+r)-f(t)|^p dt \longrightarrow 0 \quad (r\to 0)$$

証明　任意の $\varepsilon>0$ に対して，$g\in C_{00}$ が存在して，$\|f-g\|_p<\varepsilon/3$．よって，十分小さな r に対して，

$$\left\{\int |f(t+r)-f(t)|^p dt\right\}^{1/p} \leq \left\{\int |f(t+r)-g(t+r)|^p dt\right\}^{1/p}$$
$$+\left\{\int |g(t+r)-g(t)|^p dt\right\}^{1/p}+\left\{\int |g(t)-f(t)|^p dt\right\}^{1/p}<\varepsilon. \quad \blacksquare$$

定理 6.1 の証明　可測な f,g に対しては，$t-s\equiv r$ とおくことにより，$f*g=g*f$ は明らかである．（1）も明らかである．（2）を示す．

$$\|f*g\|_2^2=\|g*f\|_2^2$$
$$=\int dt\int \overline{f(s)}\,\overline{g(t-s)}ds\int f(r)g(t-r)dr$$
$$\leq\int \overline{g(t)}g(t)dt\left|\int f(s)ds\right|^2\leq\|f\|_1^2\|g\|_2^2.$$

（3）は〈補題 6.2〉より，$1/p+1/q=1$，$p\geq 1$ である p,q に対し，

$$\left|\int f(t+r-s)g(s)ds-\int f(t-s)g(s)ds\right|=\left|\int \{f(t-s+r)-f(t-s)\}g(s)ds\right|$$
$$\leq\left\{\int |f(t-s+r)-f(t-s)|^p ds\right\}^{1/p}\left\{\int |g(s)|^q ds\right\}^{1/q}$$
$$\longrightarrow 0 \quad (r\to 0). \quad \blacksquare$$

6.1.2 Banach* 代数 $L^1 \equiv L^1(\boldsymbol{R})$

L^1 はノルム $\|\cdot\|_1$ に関して複素 Banach 空間であるが，さらに有効な構造をもっている．各 $f \in L^1$ に対して，対合 $*$

$$f^*(t) = \overline{f(-t)}$$

を定義すると，このとき次の定理を得る．

〈定理 6.3〉 空間 L^1 および C_{00} は，$f*g$ を積演算とする可換代数で，さらに，$f \to f^*$ を対合とする $*$ 代数である (i.e., $(\alpha f)^* = \bar{\alpha} f^*$，$f^{**}(\equiv (f^*)^*) = f$，$(f*g)^* = g^* * f^*$，$f*(g+h) = f*g + f*h$)．

証明 $(f*g)^*(t) = \overline{(f*g)(-t)} = \int \overline{f(-t-s)}\,\overline{g(s)}\,ds$

$$= \int \overline{f(-t+s)}\,\overline{g(-s)}\,ds = \int g^*(t-s) f^*(s)\,ds = g^* * f^*(t).$$

他は明らかである．▮

〈定理 6.4〉 L^1 は Banach $*$ 代数である．すなわち，

$$\|f^*\|_1 = \|f\|_1, \qquad \|f*g\|_1 \leq \|f\|_1 \cdot \|g\|_1.$$

証明は容易である．

6.1.3 近似単位列

L^1-代数において，函数列 $\{e_n\}$ が近似単位列であるとは，$e_n^* = e_n$，$\|e_n\|_1 = 1$，$\|e_n * f - f\|_1 \to 0$ $(\forall f \in L^1)$ のときをいう．たとえば，$a_n = (n/2) 1_{[-1/n, 1/n]}$ とおくと，$\{a_n\}$ は近似単位列である．実際，$\forall f \in C_{00}$ に対して，

$$(a_n * f)(t) - f(t) = (f * a_n)(t) - f(t)$$

$$= \int f(t-s) a_n(s)\,ds - \int f(t) a_n(s)\,ds$$

$$= \frac{n}{2} \int_{-1/n}^{1/n} \{f(t-s) - f(s)\}\,ds \to 0 \quad (\text{一様収束}).$$

これを用いると，任意の $f \in L^1$ に対して，$\|a_n * f - f\|_1 \to 0$ がいえ，また $a_n^* = a_n$，$\|a_n\|_1 = 1$ は明らかである．さらに，

$$b_n = a_n * a_n = (n/2)^2 1_{[-1/n,1/n]} * 1_{[-1/n,1/n]}$$

とおくと $\{b_n\}$ も近似単位列であり，$b_n \in C_{00}^+$ である．また，任意の $f \in L^2$ に対して

$$\|b_n * f - f\|_2 \to 0.$$

本節の 6.1.5 項で近似単位列として有用な Poisson 核を説明する．

6.1.4　関数 $f \in L^1$ の Fourier 変換

$f \in L^1$ に対して，積分

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega t} f(t)\, dt$$

を f の **Fourier 変換**という．これに関して次の定理が成立する．

〈定理 6.5〉（1）　$\sup\{|\hat{f}(\omega)|;\ -\infty < \omega < \infty\} \leqq \dfrac{1}{\sqrt{2\pi}} \|f\|_1$

（2）　$\hat{f}$ は $\boldsymbol{R}$ 上の連続関数である．

（3）　**Riemann-Lebesgue**　$\hat{f} \in C_0(\boldsymbol{R})$，すなわち $\lim_{\omega \to \pm\infty} \hat{f}(\omega) = 0$.

証明　（1）は明らか．（2）は

$$|\hat{f}(\omega+h) - \hat{f}(\omega)| \leqq \frac{1}{\sqrt{2\pi}} \int |e^{-iht} - 1| \cdot |f(t)|\, dt$$

からすぐにいえる．（3）は Fourier 変換を変形して，次の式を得る．

$$\hat{f}(\omega) = -\frac{1}{\sqrt{2\pi}} \int e^{-i\omega(t+\pi/\omega)} f(t)\, dt = -\frac{1}{\sqrt{2\pi}} \int e^{-i\omega t} f\left(t - \frac{\pi}{\omega}\right) dt.$$

ゆえに，

$$2\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int e^{-i\omega t} \left\{ f(t) - f\left(t - \frac{\pi}{\omega}\right) \right\} dt$$

より，

$$\left|\hat{f}(\omega)\right| \leqq \frac{1}{2\sqrt{2\pi}} \int \left| f(t) - f\left(t - \frac{\pi}{\omega}\right) \right| dt \to 0 \quad (\omega \to \pm\infty).$$

最後の収束は，〈補題 6.2〉による．▮

〈系 6.6〉　次式が成立する．

$$\lim_{\omega \to \pm\infty} \int_{-\infty}^{\infty} f(t) \sin \omega t\, dt = \lim_{\omega \to \pm\infty} \int_{-\infty}^{\infty} f(t) \cos \omega t\, dt = 0.$$

なお，任意の $f, g \in L^1$ に対して，次の関係が成立する：

（1） $\frac{1}{\sqrt{2\pi}}(f*g)^\wedge(\omega)=\hat{f}(\omega)\hat{g}(\omega)$；$(f^*)^\wedge(\omega)=\overline{\hat{f}(\omega)}$；

（2） $f(t+a)$ と $e^{ibt}f(t)$ の Fourier 変換は各々 $e^{ia\omega}\hat{f}(\omega)$ と $\hat{f}(\omega-b)$.

これらの証明は容易である.

6.1.5 Poisson 核

Fourier 変換の具体例としては種々なものがあるが，重要な一例を述べる.
変数 $x\in\boldsymbol{R}$ の函数 $e^{-\alpha|x|}(\alpha>0)$ に対して，

$$p_\alpha(t)\equiv\left(\frac{1}{\sqrt{2\pi}}\pi e^{-\alpha|x|}\right)^\wedge(t)$$

を Poisson 核という：

$$p_\alpha(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-itx}\frac{1}{\sqrt{2\pi}}e^{-\alpha|x|}dx=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{itx}\frac{1}{\sqrt{2\pi}}e^{-\alpha|x|}dx.$$

すなわち，

$$p_\alpha(t)=\frac{2}{\sqrt{2\pi}}\int_0^\infty\frac{1}{\sqrt{2\pi}}e^{-\alpha|x|}\cos tx\,dx=\frac{1}{\pi}\frac{\alpha}{\alpha^2+t^2}\quad(\alpha>0).$$

この式から

$$\int_{-\infty}^{\infty}p_\alpha(t)\,dt=1,\quad\lim_{\alpha\to0}\int_{|t|>\delta}p_\alpha(t)\,dt=0\quad(\forall\delta>0)$$

を得，さらに $\|f*p_\alpha-f\|_1\to0\ (\alpha\to0)$ が得られる．つまり $\{p_\alpha;\alpha>0\}$ は近似単位系である．また有用な式として，任意の $f\in L^1$ に対して

$$\begin{aligned}f*p_\alpha(t)&=\frac{1}{\sqrt{2\pi}}\iint f(t-s)\frac{1}{\sqrt{2\pi}}e^{-\alpha|x|}e^{isx}\,dx\,ds\\&=\frac{1}{\sqrt{2\pi}}\iint f(s)\frac{1}{\sqrt{2\pi}}e^{-\alpha|x|}e^{itx}e^{-isx}dx\,ds\\&=\frac{1}{\sqrt{2\pi}}\int\hat{f}(x)e^{-\alpha|x|}e^{itx}dx\end{aligned}$$

を得る．なお，**Fourier 逆変換** $g\to\check{g}$ を

$$\check{g}(t)=\frac{1}{\sqrt{2\pi}}\int g(\lambda)e^{it\lambda}d\lambda\quad\text{a.e.}$$

によって定義する．とくに $f,\hat{f}\in L^1$ のとき，次が成立する：

$$\check{\hat{f}}=f$$

6.1.6　函数 $f\in L^2$ の Fourier 変換

〈定理 6.7〉　$f\in L^1\cap L^2$ に対して，$\hat{f}\in L^2$ かつ $\|f\|_2=\|\hat{f}\|_2$.

証明

$$\int\overline{f(t)}f*p_\alpha(t)dt=\frac{1}{\sqrt{2\pi}}\iint\overline{f(t)}\hat{f}(\omega)e^{-\alpha|\omega|}e^{it\omega}d\omega\,dt$$

$$=\frac{1}{\sqrt{2\pi}}\int\hat{f}(\omega)e^{-\alpha|\omega|}d\omega\overline{\int f(t)e^{-it\omega}dt}$$

$$=\int|\hat{f}(\omega)|^2e^{-\alpha|\omega|}d\omega.$$

ここで，$\alpha\to0$ のとき，左辺$\to\int|f(t)|^2dt$，右辺$\to\int|\hat{f}(\omega)|^2d\omega$. ゆえに，結論の式を得る．▮

このことを用いて $f\in L^2$ の Fourier 変換が定義される：

$$f_T(t)\equiv\begin{cases}f(t) & (|t|\leq T),\\ 0 & (|t|>T)\end{cases}$$

とおくと，$f_T\in L^1\cap L^2$ であり，$\hat{f}_T$ が定義される．したがって，変換 $f\to\hat{f}$ ($\forall f\in L^1$) の線形性を用いて，

$$\|\hat{f}_S-\hat{f}_T\|_2=\|(f_S-f_T)^\wedge\|_2=\|f_S-f_T\|_2\to0\quad(S,T\to\infty).$$

すなわち，$\{\hat{f}_T\}$ は Cauchy 列であり，L^2 の完備性から

$$\|\hat{f}_T-F\|_2\to0\quad(T\to\infty)$$

なる $F\in L^2$ が存在し，この F は函数 $f\in L^2$ によって a.e. で一意に定まる（一意性の証明は，$f\in L^1\cap L^2$ のときはすぐいえる．$f\notin L^1$，$\in L^2$ のときは，$L^1\cap L^2$ の函数によって f を $\|\cdot\|_2$-近似することによっていえる）.

このようにして，すべての $f\in L^2$ に対して一意に定まる $F\in L^2$ を，f の **Fourier** 変換といい，これを

$$\hat{f}(\omega)(=\mathcal{F}f(\omega))=\underset{T\to\infty}{\mathrm{l.i.m}}\frac{1}{\sqrt{2\pi}}\int_{-T}^{T}f(t)e^{-it\omega}dt$$

で表す (l. i. m. は "limit in the mean" の略記号).

6.1.7 Fourier 解析における基本定理

ここでは Fourier 解析において最も重要とされている 3 つの基本定理について述べる.

〈定理 6.8〉 Planchereł の定理：

（1） $\|f\|_2=\|\hat{f}\|_2=\|\mathcal{F}f\|_2,\ \forall f\in L^2.$

（2） $\langle f,g\rangle=\langle \hat{f},\hat{g}\rangle(=\langle\mathcal{F}f,\mathcal{F}g\rangle),\ \forall f,g\in L^2.$

証明 （1）は前項の後半の記述ですでに明らかであり，（2）は（1）と Hilbert 空間の内積の偏極化（付章 1 節 1.1 項参照）から明らか. ▌

作用素 $\mathcal{F}:L^2\to L^2$ は Hilbert 空間 L^2 上のユニタリー作用素であり，$\mathcal{F}$ の逆作用素 $\mathcal{F}^{-1}$ は,

$$(\mathcal{F}^{-1}\xi)(t)=\underset{T\to\infty}{\mathrm{l.i.m}}\frac{1}{\sqrt{2\pi}}\int_{-T}^{T}\xi(\omega)e^{it\omega}d\omega,\qquad \forall\xi\in L^2.$$

次に Bochner の定理の説明に移ろう. μ が $\boldsymbol{R}$ 上の有界非負な単調増加函数であるとき，μ による Riemann-Stieltjes 積分

$$\varphi(t)=\int_{-\infty}^{\infty}e^{i\omega t}d\mu(\omega) \tag{6.1}$$

を μ の **Stieltjes 変換**，または μ の**特性函数**という. このとき函数 φ は,

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\bar{\alpha}_j\alpha_k\varphi(t_j-t_k)\geq 0 \tag{6.2}$$

を満たす連続函数である. ここで $\alpha_1,\cdots,\alpha_n$ は任意有限個の複素数で，$t_1,\cdots,t_n\in\boldsymbol{R}$ である. 一般に，(6.2) を満たす函数 φ を**正の定符号函数** (p.d. と略記) という.

註 p.d. な函数 φ は次の性質をもつ：

$$\varphi(-t)=\overline{\varphi(t)},\qquad |\varphi(t)|\leq\varphi(0),\qquad \forall t\in\boldsymbol{R}.$$

これは，代数的な計算によって示されるが，その証明は略する.

〈定理 9.9〉 Bochner の定理： $\boldsymbol{R}$ 上の可測な正の定符号函数 φ に対して，一意に $\boldsymbol{R}$ 上の有界非負な単調増加函数 μ が存在して (6.1) が満たされる.

この定理の証明は，それほど容易ではないが，ここでは略する．たとえば，文献［G. 3, R. 2］を参照．

註　有界非負単調増加函数 μ と $\boldsymbol{R}$ 上の有限正値測度が1対1に対応している．さらに，φ が $\varphi(0)=1$ を満たすと，これに対応する測度は $\boldsymbol{R}$ 上の確率測度になる．したがって，Bochner の定理は函数解析のみならず確率論においても非常に重要な役割を果たしている．

最後に，Stone の定理を述べる．$\mathcal{H}$ を与えられた Hilbert 空間とし，$\{U_t; -\infty<t<\infty\}$ を $\mathcal{H}$ 上のユニタリー作用素の1径数群で，パラメーター t に関して連続であるとする．すなわち，$U_{s+t}=U_sU_t, U_{-t}=U_t^*$ であり，$t\longmapsto\langle x, U_t x\rangle$ $(\forall x\in\mathcal{H})$ が $\boldsymbol{R}$ 上の連続函数とする．このとき，

〈定理 6.10〉 **Stone の定理**：$\boldsymbol{R}$ 上の単位の分解 $\{E_\lambda; -\infty<\lambda<\infty\}$ が存在して，

$$U_t=\int e^{it\lambda}dE_\lambda, \qquad \forall t\in\boldsymbol{R} \tag{6.3}$$

が成立する．ここで積分（6.3）は Fourier-Stieltjes 積分

$$\langle x, U_t y\rangle=\int e^{it\lambda}d\langle x, E_\lambda y\rangle, \qquad \forall x, y\in\mathcal{H}$$

の意味である．

この節の内容は文献［G. 3, K. 5, N. 2, U. 15］などに詳しい．

6.2　信号解析

信号とは光・音・色・形など，あるいはこれらの組み合わせを用いて，意志を伝達する方法をいう．これら光・音などを一種の記号であると考え，信号はこれら記号の系列によって形成されると考えられる．通常，信号を形成する記号は実数，あるいは複素数で表される．しかしながら，時によっては数によって表示されない場合もあるが，そのような場合も何らかの仕方で数に換えて論ずることが多い．記号の系列は離散的な形 $\{x_n; n=0, \pm1, \pm2, \cdots\}$，$\{x_n; n=0,$

1, 2, ⋯} や連続的な形 $\{x_t; 0 \leq t < \infty\}$, $\{x_t; -\infty < a \leq t \leq b \leq +\infty\}$ の 2 つの場合が考えられ，各々単位時間ごとに，あるいは各時刻 t で 0 または 1，有限個の数字の集合の元，任意の実数または複素数などをとると考えられる．これらは光・音波・電信（＝電気的な信号の記号列），あるいはある種の物体の運動などの観測の結果の量的表示と考えるのが自然である．つまり，信号とは観測の対象となる物理的な系に関しての測定結果であり，数値による（記号系列としての）物理的表現形式としてとらえられたものであり，一般には，波形としてグラフ表示される（信号波形）．

上述のように，信号とはスカラー値な時間函数であり，この函数を信号函数という．以後，$-\infty < t < \infty$ のように，時刻 t が数直線 $\boldsymbol{R}$ 全体にわたる場合を主として考える．他の場合は $\boldsymbol{R}$ の部分区間とか自然数のみに限定する特別な場合である．

一般に，信号 $x_t = x(t)$ は決定論的なものとランダムなものに大別される．つまり，$t = t_0$ における信号の表示が $x(t_0)$ として確定される場合と，ランダムさが伴って確定的でない場合の 2 者に分けられる．前者の場合は，信号 $x(\cdot)$ は Hilbert 空間 $L^2(-\infty, \infty)$ や Banach 空間 $L^\infty(-\infty, \infty)$ に属する函数として，後者の場合は二次のモーメントをもつランダム過程（確率過程）として論ぜられる．

上で説明した信号は，実数直線上で定義された何らかの可測性をもつ函数として表示される．それらの基本的な信号函数を与えることにする．

（1） **正弦波信号**

$$
\begin{aligned}
x(t) &= \mathrm{Re}(\exp(a + i(\theta + 2\pi f t))) \\
&= \exp a \cdot \mathrm{Re}(\exp(i(\theta + 2\pi f t))) = e^a \mathrm{Re}(e^{i(\theta + \omega t)}) \\
&= e^a \cos(\theta + \omega t).
\end{aligned}
$$

ここで $\exp a$ が振幅，θ が位相，f が周波数（>0），$\omega = 2\pi f$ であり，これらはすべて実数（i は虚数単位）である．

（2） **周期的信号**

$$
x(t) = x(t + T), \qquad -\infty < t < \infty
$$

を満たす最小の $T>0$ を $x(t)$ の周期という．このように周期をもつ信号を周期的信号という．（1）の正弦波信号では周期 $T=1/f$ である．

（3）　**有界振幅信号**（または**振幅制限信号**）

$$|x(t)|\leqq \text{const. a.e.}$$

である可測函数 $x=x(t)$ を有界振幅信号という．このような $x(\cdot)$ の全体は $L^\infty(\boldsymbol{R})$ で表される．これはノルム $\|x\|_\infty\equiv\text{ess sup}\{|x(t)|;t\in\boldsymbol{R}\}=\inf\{M>0;|x(t)|\leqq M\text{ a.e.}\}$ に関して Banach 空間である．

（4）　**エネルギー有限な信号**（または**有限パワー信号**）

$\boldsymbol{R}$ 上の可測函数 x で

$$\|x\|_2=\left(\int_{-\infty}^{\infty}|x(t)|^2dt\right)^{1/2}<+\infty$$

である $x(t)$ をエネルギー有限な信号といい，この信号の全体を $L^2(\boldsymbol{R})$ で表す．各 $x\in L^2(\boldsymbol{R})$ は電気信号と考え，とくに時刻 t において電信発生器が1オームの抵抗をもっていたとき流れる電流の値が信号値 $x(t)$ で与えられると考えると，ノルムの平方 $\|x\|_2^2$ は電源から抵抗に供給される**全エネルギー**に等しい．

（5）　**時間制限信号**（time limited signal）

有限区間 $[a,b]$ が存在して

$$x(t)=0\quad(t<a\text{ or }t>b)$$

である Lebesgue 可測な函数 $x(t)$ を時間制限信号という．

（6）　**帯域制限信号**（band limited signal）

定数 $W>0$ が存在して，$|\omega|\geqq W$ に対して，

$$\hat{x}(\omega)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-i\omega t}x(t)\,dt=0$$

である Lebesgue 可測な函数 $x(t)$ を帯域制限信号という．ここで，W を帯域幅，$[-\pi W,\pi W]$ を $\hat{x}$ の**周波数区間**という．一般に信号 $x\in L^2(\boldsymbol{R})$ に対して，Fourier 変換 $\hat{x}$ を計算し $|\hat{x}(\omega)|^2$ を信号 x の**エネルギー密度**（またはパワースペクトラム）という．

標準的な帯域制限信号として，次の正弦波信号（i），(ii）が用いられる．

（i） **標本函数** (sampling function)

$$\varphi(t)=\frac{W}{\pi}\frac{\sin Wt}{Wt}, \qquad \varphi(0)=\frac{W}{\pi}.$$

ここで

$$\frac{\sin 0}{0}=1$$

とした．

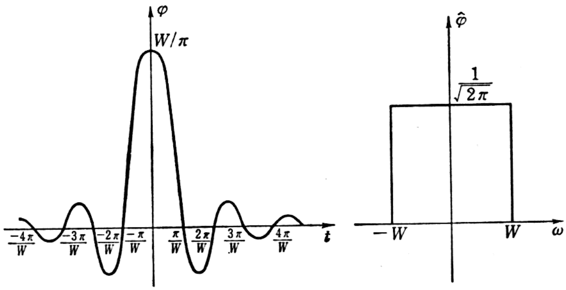


図 6.1

（ii） **標本函数系**

標本函数 $\varphi(\cdot)$ の時刻 $t-n\pi/W$ における信号値に $\sqrt{\dfrac{\pi}{W}}$ を乗じたものを

$$\varphi_n(t)\equiv\sqrt{\frac{\pi}{W}}\,\varphi(t-n\pi/W)=\sqrt{\frac{W}{\pi}}\frac{\sin W(t-n\pi/W)}{W(t-n\pi/W)},$$

$$n=0,\ \pm 1,\ \pm 2,\cdots$$

で表す．このとき $\{\varphi_n\}$ を**標本函数系**という．

ここで，標本函数の Fourier 変換

$$\hat{\varphi}(\omega)=\underset{T\to\infty}{\mathrm{l.i.m}}\frac{1}{\sqrt{2\pi}}\int_{-T}^{T}\frac{W}{\pi}\frac{\sin Wt}{Wt}e^{-i\omega t}dt=\begin{cases}\dfrac{1}{\sqrt{2\pi}} & (|\omega|\leq W),\\ 0 & (|\omega|>W)\end{cases}$$

を使うと，

$$\hat{\varphi}_n(\omega)=\sqrt{\frac{\pi}{W}}\,e^{-i\omega n\pi/W}\hat{\varphi}(\omega), \qquad n=0,\ \pm 1,\ \pm 2,\cdots$$

と表せる．

〈**定理 6.11**〉 標本函数系 $\{\varphi_n\}$ は Hilbert 空間 $L^2(\boldsymbol{R})$ において正規直交系（ONS）をなす．

証明　Plancherel の定理より

$$\int_{-\infty}^{\infty}\overline{\varphi_m(t)}\varphi_n(t)dt=\int_{-\infty}^{\infty}\overline{\hat{\varphi}_m(\omega)}\hat{\varphi}_n(\omega)d\omega$$

$$=\frac{1}{2\pi}\int_{-W}^{W}\frac{\pi}{W}e^{i\omega(m-n)\pi/W}d\omega$$

$$=\frac{1}{2W}\frac{\exp(iW(m-n)\pi/W)-\exp(-iW(m-n)\pi/W)}{i(m-n)\pi/W}$$

$$=\frac{1}{2W}\frac{2\sin\{(m-n)\pi\}}{(m-n)\pi/W}$$

$$=\frac{\sin((m-n)\pi)}{(m-n)\pi}=\delta_{m,n}.\ \blacksquare$$

信号函数は，一般に連続な時間軸 $\boldsymbol{R}$ 上で定義されたスカラー値な函数で，数学的にはある種の可測性をもった時間変動な函数であると考えられてきたが，このような時間が連続的に変化する信号を離散的な信号に変換することを**標本化**という．

たとえば，連続的に動く被写体を映写機によって映す場合，単位時間に一定のコマ数が映され，それによって各瞬間ごとに被写体の映像が結ばれ，動く被写体が再現される．この，単位時間に必要なコマ数を割り出すための理論構成が次に述べる Shanonn の**標本化定理**である．

〈**定理 6.12**〉 帯域制限信号で帯域幅が W である $x\in L^2(\boldsymbol{R})$ について，次の展開式が成立する．

$$x(t)=\sum_{n=-\infty}^{\infty}\alpha_n\cdot\sqrt{\frac{\pi}{W}}\varphi_n(t).$$

ここで，$\alpha_n=x(n\pi/W)$ であり，$\{\varphi_n(t)\}$ は標本函数系である（この α_n を時刻 $t=n\pi/W$ における信号の標本値という）．

証明

$$x(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\hat{x}(\omega)e^{i\omega t}d\omega=\frac{1}{\sqrt{2\pi}}\int_{-W}^{W}\hat{x}(\omega)e^{i\omega t}d\omega.$$

これは $C_0(\boldsymbol{R})$ に属する函数である．一方，$\hat{x}(\cdot)$ の Fourier 級数展開は，

$$\hat{x}(\omega)=\sum_{n=-\infty}^{\infty}c_n\exp(-in\pi\omega/W),$$

$$c_n=\frac{1}{2W}\int_{-W}^{W}e^{in\pi\omega/W}\hat{x}(\omega)d\omega=\frac{\sqrt{2\pi}}{2W}x\left(\frac{n\pi}{W}\right).$$

これら $\hat{x}(\omega)$ と c_n を最初の等式の右辺に代入すると，

$$\begin{aligned}
x(t)&=\frac{1}{\sqrt{2\pi}}\sum_{n=-\infty}^{\infty}c_n\int_{-W}^{W}\exp(-in\pi\omega/W)\exp(i\omega t)d\omega\\
&=\frac{1}{\sqrt{2\pi}}\sum\frac{\sqrt{2\pi}}{2W}x\left(\frac{n\pi}{W}\right)\int_{-W}^{W}e^{i\omega(t-n\pi/W)}d\omega\\
&=\frac{1}{2W}\sum x\left(\frac{n\pi}{W}\right)\frac{\exp(iW(t-n\pi/W))-\exp(-iW(t-n\pi/W))}{i(t-n\pi/W)}\\
&=\frac{1}{2W}\sum x\left(\frac{n\pi}{W}\right)\frac{2\sin(W(t-n\pi/W))}{t-n\pi/W}\\
&=\sum x\left(\frac{n\pi}{W}\right)\frac{\sin(W(t-n\pi/W))}{W(t-n\pi/W)}\\
&=\sum x\left(\frac{n\pi}{W}\right)\sqrt{\frac{\pi}{W}}\varphi_n(t). \quad \blacksquare
\end{aligned}$$

この標本化定理は次のことを意味している：“帯域制限信号 x（$\hat{x}$ が πW 以上の周波数成分を含まない）において，帯域幅 W によって指定される各離散的な時点における信号の標本値が得られれば，すべての時点 t における信号値が一意に定まる”．この定理は，数学的には，周波数区間 $[-\pi W, \pi W]$ に含まれるエネルギー有限な信号の全体は標本函数系 $\{\varphi_n\}$ によって張られる Hilbert 空間と一致し，その空間の各ベクトル（各信号）を直交系 $\{\varphi_n\}$ により Fourier 展開したものが標本化の終結の展開式であることを示している．要するに，標本化定理とはひとつの特殊な ONS による直交展開なのである．

この節の議論には，文献 [H. 7, S.1, T.3] などが参考になる．

6.3 Karhunen-Loéve 展開

前節で議論した標本化定理はランダム性をもたない信号の直交展開に関するものであったが，ここではランダム性をもつ信号についての直交展開を論ずる．ここで扱う信号函数は，これまでと同様に一般性をもたせるため複素数値をとるものとする．

本論に入る前に，積分作用素について必要な部分の説明を行う．ここで問題とする空間は，有限区間 $[a,b]$ 上の Hilbert 空間 $L^2(a,b)$ である．正方形 $[a,b]^2=[a,b]\times[a,b]$ 上の Lebesgue 可測函数 $K(s,t)$（つまり2変数函数として同時可測であるもの）に対して，積分値

$$(Kx)(s)=\int_a^b K(s,t)x(t)dt, \qquad \forall x\in L^2(a,b) \tag{6.4}$$

がつねに存在し，しかもこの値 $(Kx)(s)$ が $s\in[a,b]$ の函数として可測で $L^2(a,b)$ に属するとき，写像 $K(x\in L^2(a,b)\to Kx\in L^2(a,b))$ を**積分作用素**といい，与えられた函数 $K(\cdot,\cdot)$ を**核函数**という．この作用素が意味をもち，効果的な役割を演ずるのは K が Schmidt 作用素（i.e., $K\in \boldsymbol{S}(L^2(a,b))$，付章4節4.6項参照）の場合である．これに関する重要な定理を述べよう．

〈**定理 6.13**〉 $[a,b]^2$ 上の Lebesgue 可測函数 $K(\cdot,\cdot)$ に対する積分作用素が $K\in \boldsymbol{S}(L^2(a,b))$ であるための必要十分条件は，$K(\cdot,\cdot)\in L^2([a,b]^2)$ となることである．また，次の（1），（2）が成り立つ（a.e. の意味で）．

（1）　$K(s,t)=\overline{K(t,s)} \Longleftrightarrow K=K^*$（自己共役），

（2）　$K(s,t)$ が正の定符号 $\Longleftrightarrow K$ が正の作用素．

さらに，対応 $K(\cdot,\cdot)\in L^2([a,b]^2)\longleftrightarrow K\in \boldsymbol{S}(L^2(a,b))$ は1対1線形で

$$\left(\int_a^b\int_a^b |K(s,t)|^2 ds\,dt\right)^{1/2}=\|K\|_2 \quad (=\text{Schmidt ノルム})$$

が成立し，$L^2([a,b]^2)$ と $\boldsymbol{S}(L^2(a,b))$ は Hilbert 空間として同一視できる．

この定理によって，作用素 K の固有値展開 $\sum \lambda_k x_k\otimes\bar{x}_k$ が可能になる（付

章 4 節 4.6 項)．ここで，この展開式に関する重要な Mercer の定理を Hilbert 空間 $L^2(a,b)$ 上で論じよう．

〈**定理 6.14**〉 **Mercer の定理**：$L^2(a,b)$ 上の正の積分作用素

$$K=\sum_{k=1}^{\infty}\lambda_k x_k\otimes\bar{x}_k\in \boldsymbol{S}(L^2(a,b))$$

の核函数 $K(s,t)$ が正方形 $[a,b]^2$ 上で連続（同時連続という）であるとき，各固有函数 $x_k(\cdot)$ は $[a,b]$ 上で連続であり

$$K(s,t)=\sum_{k=1}^{\infty}\lambda_k x_k(s)\overline{x_k(t)},\qquad \forall s,t\in[a,b] \tag{6.5}$$

と表される．さらに，次の（1），（2）が成立する．

（1） 等式 (6.5) の右辺は絶対収束し，$[a,b]^2$ 上での一様収束であり，かつ $L^2([a,b]^2)$ でノルム収束する．すなわち，$n\to\infty$ のとき

$$\sup_{s,t\in[a,b]}\left|K(s,t)-\sum_{k=1}^{n}\lambda_k x_k(s)\overline{x_k(t)}\right|\to 0, \tag{6.6}$$

かつ

$$\int_a^b\int_a^b\left|K(s,t)-\sum_{k=1}^{n}\lambda_k x_k(s)\overline{x_k(t)}\right|^2 ds\ dt\to 0. \tag{6.7}$$

（2） $K\in \boldsymbol{T}(L^2(a,b))$ であり

$$\int_a^b K(t,t)dt=\sum_{k=1}^{\infty}\lambda_k=\mathrm{tr}(K).$$

証明 固定した各 $s\in[a,b]$ に対して $K(s,\cdot)\in L^2(a,b)$ であるから，付章の〈定理 4.7〉の（3）において A に K を，y に $K(s,\cdot)$を代入すると，s に従属してスカラー $\alpha_k(s)=\langle x_k,K(s,\cdot)\rangle=\lambda_k x_k(s)$ および t の函数 $\beta(s,t)\in L^2(a,b)$ が存在して

$$K(s,\cdot)=\sum_{k=1}^{\infty}\alpha_k(s)\overline{x_k(\cdot)}+\beta(s,\cdot)$$

と表される．ここで再び付章の〈定理 4.7〉によって $\beta(s,\cdot)\perp x_k(k=1,2,\cdots)$，さらに，$K\beta(s,\cdot)=0$ がいえ，a.e. $r\in[a,b]$ に対して

$$0=K\beta(s,\cdot)(r)=\int_a^b K(r,t)\beta(s,t)dt=\int_a^b K(t,r)\overline{\beta(s,t)}dt.$$

したがって，$K(\cdot,\cdot)$ の連続性から

$$0=\int_a^b K(t,s)\overline{\beta(s,t)}dt=\int_a^b \sum\lambda_k\overline{x_k(s)}x_k(t)\overline{\beta(s,t)}dt+\int_a^b|\beta(s,t)|^2dt$$
$$=\sum_{k=1}^{\infty}\lambda_k x_k(s)\langle\beta(s,\cdot),x_k\rangle+\int_a^b|\beta(s,t)|^2dt$$
$$=\int_a^b|\beta(s,t)|^2dt.$$

ゆえに，$\beta(s,t)=0$ a.e. $t\in[a,b]$，すなわち

$$K(s,\cdot)=\sum_{k=1}^{\infty}\lambda_k x_k(s)\overline{x_k(\cdot)},\qquad s\in[a,b]\quad(\textstyle\sum\text{ は }L^2\text{-収束}).$$

(1)，(2) については3段階に分けて証明する．

［第1ステップ］　函数 $R_n(s,t)\equiv K(s,t)-\sum_{k=1}^{n}\lambda_k x_k(s)\overline{x_k(t)}$ が同時連続なことは明らかである．$R_n(t,t)\geq 0$ ($\forall t\in[a,b]$) であることを示す．証明の第一段目の結論から

$$R_n(s,t)=\sum_{k=n+1}^{\infty}\lambda_k x_k(s)\overline{x_k(t)}\quad(\textstyle\sum\text{ は各変数ごとの }L^2\text{-収束}).$$

函数 $R_n(\cdot,\cdot)$ に対応する作用素 ((6.4) 参照) を R_n と書くと，スペクトル表示で $R_n=\sum_{k=n+1}^{\infty}\lambda_k x_k\otimes\bar{x}_k$ と書けることから

$$\langle R_n y,y\rangle=\sum_{k=n+1}^{\infty}\lambda_k|\langle y,x_k\rangle|^2\geqq 0,\qquad y\in L^2(a,b).\tag{6.8}$$

これより $R_n(t,t)\geq 0$ ($\forall t\in[a,b]$) が容易にいえる．

［第2ステップ］　函数項級数 $\sum\lambda_n x_n(s)\overline{x_n(t)}$ は絶対収束し，各変数ごとに一様収束することを示す．［第1ステップ］によって

$$R_n(t,t)=K(t,t)-\sum_{k=1}^{n}\lambda_k|x_k(t)|^2\geq 0,\qquad n=1,2,\cdots.$$

ゆえに，

$$\sum_{k=1}^{\infty}\lambda_k|x_k(t)|^2\leq K(t,t)\leq M<\infty.\tag{6.9}$$

したがって，

$$\left|\sum_{k=n}^{m}\lambda_k x_k(s)\overline{x_k(t)}\right|^2=\left|\sum_{k=n}^{m}\sqrt{\lambda_k}\ \overline{x_k(s)}\sqrt{\lambda_k}x_k(t)\right|^2$$

$$\leq\left(\sum_{k=n}^{m}\lambda_k|x_k(s)|^2\right)\left(\sum_{k=n}^{m}\lambda_k|x_k(t)|^2\right)$$

$$\leq M\sum_{k=n}^{m}\lambda_k|x_k(s)|^2\to 0\quad (m\geqq n\to\infty).$$

ゆえに，収束は各固定した s で t に関して一様収束である．絶対収束することは明らかである．

［第3ステップ］ $K(s,t)=\sum_{k=1}^{\infty}\lambda_k x_k(s)\overline{x_k(t)}$, $\forall s,t\in[a,b]$．実際，級数 $\sum=\sum\lambda_k x_k(s)\overline{x_k(t)}$ は絶対収束するので，その総和を $K_0(s,t)$ とおく．［第2ステップ］によって，これは t の連続函数である．また，固定した s に対して $K(s,t)=$ 級数 $\sum$ の L^2-極限．したがって $K(s,t)=K_0(s,t)$ a.e. $t\in[a,b]$．また $K(s,t)$, $K_0(s,t)$ がともに t の連続函数であることを用いて，結局 $K(s,t)=K_0(s,t)$ がすべての $s,t\in[a,b]$ で成立する．

［第4ステップ］ $S_n(t)\equiv\sum_{k=1}^{n}\lambda_k|x_k(t)|^2\in C[a,b]$ は，$S_n(t)\uparrow K(t,t)$ であり，Dini の単調函数列の一様収束定理によって，一様収束する．これから［第2ステップ］の (6.9) を用いて，一様収束の結論 (6.6) がいえ，同時に (6.7) を得る．

（2）は容易である：$K(t,t)=\sum\lambda_k|x_k(t)|^2$（一様収束）．ゆえに，

$$\int_a^b K(t,t)dt=\sum\lambda_k\int_a^b|x_k(t)|^2dt=\sum\lambda_k=\mathrm{tr}\,K. \blacksquare$$

Mercer の定理の適応例として，しばしば登場する函数

$$K(s,t)=\min(s,t)\equiv s\wedge t,\qquad s,t\in[0,T]$$

について調べよう．積分方程式

$$\int_0^T(s\wedge t)x(t)dt=\lambda x(s),\qquad 0\leq s\leq T \tag{6.10}$$

を考える．当然 x は Lebesgue 可積分性が前提となる．これは

$$\int_0^s tx(t)dt+\int_s^T sx(t)dt=\lambda x(s)$$

とかき直せる．これから $x(\cdot)$ は微分可能であり，微分すると

$$\int_s^T x(t)dt=\lambda x'(s),\qquad 0<s<T. \tag{6.11}$$

これもまた微分可能で，両辺を微分すると

$$-x(s)=\lambda x''(s),\qquad 0<s<T. \tag{6.12}$$

(6.10) と (6.11) および境界条件 $x(0)=x'(T)=0$ を満たすことから，微分方程式 (6.12) の解は

$$x(t)=a\sin(t/\sqrt{\lambda})$$

となり，$0=x'(T)=(a/\sqrt{\lambda})\cos(T/\sqrt{\lambda})$. これを解くと，固有値 λ_n は

$$\lambda_n=(2T/(2n+1)\pi)^2,\qquad n=0,1,2,\cdots$$

であり，対応する固有函数 x_n は

$$x_n(t)=\sqrt{\frac{2}{T}}\sin\left(\frac{(2n+1)\pi}{2T}t\right),\qquad n=0,1,2,\cdots$$

となり，これらを Mercer の定理に適用すると，

$$s\wedge t=\frac{2}{T}\sum_{n=0}^{\infty}\left(\frac{2T}{(2n+1)\pi}\right)^2\sin\left(\frac{(2n+1)\pi}{2T}s\right)\sin\left(\frac{(2n+1)\pi}{2T}t\right).$$

ここで，無限和 $\sum$ は $[0,T]^2$ 上で一様収束極限である.

註 Mercer の定理は 2 つの特徴がある．まず，$C([a,b]^2)=C[a,b]\otimes C[a,b]$（テンソル積）から，$K(\cdot,\cdot)\in C([a,b]^2)$ は $\sum f_k^{(n)}(s)g_k^{(n)}(t)$ ($\sum$ は有限和) の形の 2 変数函数の列によって一様近似される ($n\to\infty$ のとき) ことは明らかであるが，Mercer の定理は K を函数級数で近似するのではなく，それを，積分作用素 K の固有函数で直接表すという点に特徴がある．他の特徴は，作用素 K の固有値展開そのものが一様収束によって表示されていることである．

この節の後半として，目標の **Karhunen-Loéve (K-L と略記) 展開**について論ずる.

以下，連続パラメーター t ($a\leq t\leq b$) をもち，二次のモーメントを有するランダム信号 $\xi(t)=\xi(t,\cdot)$ を考える．これは与えられたある確率空間 $(\Omega,\mathcal{F},P)$ 上の確率変数系で $\xi(t,\cdot)\in L^2(\Omega)\,(\forall t\in[a,b])$ を満たす $\xi(t)$ である．$\xi(t)$ の時間差の 2 乗平均値をとり

$$E(|\xi(t+h)-\xi(t)|^2)\to 0\quad(h\to 0)$$

のとき $\xi(t)$ を **(2 乗平均) 連続**という．さらに

$$R_\xi(s,t)=E(\overline{\xi(s,\cdot)}\xi(t,\cdot)),\qquad s,t\in[a,b]$$

が存在する．これを信号 $\xi(t)$ の**相関函数**という．これが正の定符号性

$$\sum \bar{\alpha}_i \alpha_j R_\xi(t_i, t_j) \geqq 0 \quad (\textstyle\sum \text{ は有限和})$$

をもつことは明らか．まず次の定理を得る．

〈定理 6.15〉 ランダム信号 $\xi(t)$ に対して，次の性質は同値である．

（1） $\xi(t)$ は連続である．

（2） $R_\xi(s,t)$ はすべての $(s,t)\in[a,b]^2$ で連続，すなわち，2変数として同時連続である．

証明 （1）⇨（2）：$\xi(t)$ の連続性から $E(|\xi(t+k)|^2)\to E(|\xi(t)|^2)$ $(k\to 0)$ であり，これを用いて

$$\begin{aligned}
|R_\xi(s+h,t+k)-R_\xi(s,t)| &= |E(\overline{\xi(s+h)}\xi(t+k))-E(\overline{\xi(s)}\xi(t))| \\
&\leq |E((\overline{\xi(s+h)}-\overline{\xi(s)})\xi(t+k))|+|E(\overline{\xi(s)}(\xi(t+k)-\xi(t)))| \\
&\leq [E(|\xi(s+h)-\xi(s)|^2)E(|\xi(t+k)|^2)]^{1/2} \\
&\quad +[E(|\xi(s)|^2)E(|\xi(t+k)-\xi(t)|^2)]^{1/2} \\
&\to 0 \quad (h,k\to 0).
\end{aligned}$$

（2）⇨（1）：

$$\begin{aligned}
E(|\xi(t+h)-\xi(t)|^2) &= R_\xi(t+h,t+h)-R_\xi(t+h,t)-R_\xi(t,t+h)+R_\xi(t,t) \\
&= (R_\xi(t+h,t+h)-R_\xi(t,t))-(R_\xi(t+h,t)-R_\xi(t,t)) \\
&\quad -(R_\xi(t,t+h)-R_\xi(t,t)) \\
&\to 0 \quad (h\to 0). \quad \blacksquare
\end{aligned}$$

このことより，連続なランダム信号 $\xi(t)$ が与えられたとき，その相関函数 $R_\xi(\cdot,\cdot)$ は同時連続となり，対応する積分作用素 R_ξ は正値で，$R_\xi\in \boldsymbol{S}(L^2(a,b))$ であるから

$$R_\xi=\sum \lambda_k r_k \otimes \bar{r}_k, \qquad r_k\in L^2(a,b), \qquad \lambda_k \downarrow 0 \tag{6.13}$$

で表示されることが知られる．したがって，Mercer の定理が適用され，$R_\xi\in \boldsymbol{T}(L^2(a,b))$ となる．ここで λ_k と r_k は，それぞれ積分作用素 R_ξ の固有値，固有函数である：$R_\xi r_k=\lambda_k r_k$（容易に $r_k\in C[a,b]$ がいえる）．この記号のもとで

〈定理 6.16〉 **K-L 展開定理**：連続なランダム信号 $\xi(t)$ に対して，確率変数列

（1）　$\xi_k(\cdot)=\int_a^b \xi(t,\cdot)r_k(t)dt/\sqrt{\lambda_k}, \qquad k=1,2,\cdots$

は Hilbert 空間 $L^2(\Omega)$ で ONS をなし，

（2）　$\lim_{n\to\infty} E\left(\left|\xi(t,\cdot)-\sum_{k=1}^{n}\sqrt{\lambda_k}r_k(t)\xi_k(\cdot)\right|^2\right)=0.$

ここで，式（1）の右辺は $L^2(\Omega)$-値 Bochner 積分であり，（2）の lim は区間 $[a,b]$ 上の一様収束による極限である.

証明　$E(\overline{\xi(t,\cdot)}\eta(\cdot))$ $(\forall\eta\in L^2(\Omega))$ と $E(|\xi(t,\cdot)|^2)$ は区間 $[a,b]$ 上の連続函数であることから，$\xi(t,\cdot)x(t)$ $(\forall x\in L^2(a,b))$ は $L^2(\Omega)$-値 Bochner 可積分である（付章5節参照）．ゆえに，（1）の右辺は存在する．この $\{\xi_k(\cdot)\}$ が ONS であることを証明する.

$$
\begin{aligned}
\langle\xi_i,\xi_j\rangle_\Omega=E(\bar{\xi}_i\cdot\xi_j)&=\frac{1}{\sqrt{\lambda_i\lambda_j}}E\left(\int_a^b\int_a^b\overline{\xi(s,\cdot)}\overline{r_i(s)}\xi(t,\cdot)r_j(t)ds\,dt\right)\\
&=\frac{1}{\sqrt{\lambda_i\lambda_j}}\int_a^b\int_a^b E(\overline{\xi(s,\cdot)}\xi(t,\cdot))\overline{r_i(s)}r_j(t)ds\,dt\\
&=\frac{1}{\sqrt{\lambda_i\lambda_j}}\int_a^b\left(\int_a^b R_\xi(s,t)\overline{r_i(s)}ds\right)r_j(t)dt\\
&=\frac{\lambda_i}{\sqrt{\lambda_i\lambda_j}}\int_a^b\overline{r_i(t)}r_j(t)dt=\delta(i,j).
\end{aligned}
$$

次に定理の後半を証明する.

$$
\begin{aligned}
0&\le E\left(\left|\xi(t,\cdot)-\sum_{k=1}^{n}\sqrt{\lambda_k}r_k(t)\xi_k(\cdot)\right|^2\right)\\
&=E(|\xi(t,\cdot)|^2-2\sum_{k=1}^{n}\sqrt{\lambda_k}\mathrm{Re}(E(\overline{\xi(t,\cdot)}\xi_k(\cdot)\overline{r_k(t)}))+\sum_{k=1}^{n}\lambda_k|r_k(t)|^2\\
&=R_\xi(t,t)-2\sum_{k=1}^{n}\mathrm{Re}\left(\int_a^b E(\overline{\xi(t,\cdot)}\xi(s,\cdot)r_k(s)\overline{r_k(t)})ds\right)\\
&\qquad+\sum_{k=1}^{n}\lambda_k|r_k(t)|^2\\
&=R_\xi(t,t)-2\sum\mathrm{Re}\left(\int_a^b\overline{R_\xi(s,t)}r_k(s)ds\cdot\overline{r_k(t)}\right)+\sum\lambda_k|r_k(t)|^2
\end{aligned}
$$

$$=R(t,t)-2\sum\lambda_k|r_k(t)|^2+\sum\lambda_k|r_k(t)|^2$$

$$=R(t,t)-\sum_{k=1}^{n}\lambda_k|r_k(t)|^2\to 0 \quad (n\to\infty).$$

これが一様収束であることは，Mercer の定理による．▌

註 与えられたランダム信号 $\xi(t)=\xi(t,\cdot)$ に対して，〈定理 6.16〉の式（1）における2つの函数列 $\{r_k\}\subset L^2(a,b)$ と $\{\xi_k\}\subset L^2(\Omega)$ がともに ONS であるとき，定理の結論の式（2）が成立するならば展開式（6.13）が成立する．実際，任意の $s,t\in[a,b]$ に対して

$$\begin{aligned}R_\xi(s,t)&=E(\overline{\xi(s,\cdot)}\cdot\xi(t,\cdot))\\&=\lim_{m,n\to\infty}E\Bigl(\sum_{i=1}^{m}\sum_{j=1}^{n}\sqrt{\lambda_i}\ r_i(s)\overline{\xi_i(\cdot)}\sqrt{\lambda_j}\ \overline{r_j(t)}\xi_j(\cdot)\Bigr)\\&=\lim_{m,n\to\infty}\sum_{i=1}^{m}\sum_{j=1}^{n}\sqrt{\lambda_i}\ r_i(s)\sqrt{\lambda_j}\ \overline{r_j(t)}E(\overline{\xi_i(\cdot)}\xi_j(\cdot))\\&=\lim_{n\to\infty}\sum_{k=1}^{n}\lambda_k r_k(s)\overline{r_k(t)}.\end{aligned}$$

ここで lim は一様収束である．これから容易に

$$\int_a^b R_\xi(s,t)r_k(t)dt=\lambda_k r_k(s), \qquad k=1,2,\cdots.$$

したがって，展開式（6.13）が成立する．

この節の話題は主に [A.9, H.6, N.6] によっている．

6.4 不確定性関係

量子力学系は，それが m 個の自由度をもつとき，通常微分作用素＝Hamiltonian H と，それが作用する Hilbert 空間 $L^2(\boldsymbol{R}^m)$ とによって記述されている．ここで，Hamiltonian H は，m 個の座標と，それらに共役な運動量を変数とする Hamilton 函数を量子化することよって得られる．この系の状態は波動函数 $x=x(q_1,\cdots,q_m)\in L^2(\boldsymbol{R}^m)$，$\|x\|=1$ によって表され，この波動函数 x は Hamiltonian H による Schrödinger の波動方程式 $Hx=Ex$（E は H の固有値）の解として与えられる．このように量子系の議論は函数解析をベースとして展開されるのである [N.3, V.1]．

一般に，複素 Hilbert 空間 $\mathcal{H}$ と，各ベクトル $x\in\mathcal{H}$，$\|x\|=1$ に対して1次元空間 $\{\lambda x\}$ 上への射影作用素 $|x\rangle\langle x|$（あるいは $x\otimes\bar{x}$）を純粋状態と名

付ける．また，$\mathscr{H}$ の ONS $\{x_n\}$ と非負な実数列 $\{\lambda_n\}$，$\sum\lambda_n=1$，に対して，作用素

$$\rho=\sum\lambda_n|x_n\rangle\langle x_n| \quad (\textstyle\sum \text{は強収束の意味}) \tag{6.14}$$

は，正のトレース・クラス作用素で $\mathrm{tr}\,\rho=1$ である（i.e., $\rho\in\boldsymbol{T}(\mathscr{H})_{+1}$）．この ρ を系の**混合状態**あるいは**密度作用素**とよぶ．

物理量は $\mathscr{H}$ 上の必ずしも有界でない Hermite 作用素 $A, B, \cdots$ などによって表される．たとえば $\mathscr{H}=L^2(\boldsymbol{R}^m)$ の場合，

$$P_k=-i\frac{\partial}{\partial q_k} \quad (\textbf{運動量作用素}), \qquad Q_k=q_k\cdot \quad (\textbf{位置作用素}) \tag{6.15}$$

や，Hamiltonian H などはもちろん，純粋状態 $|x\rangle\langle x|$ や密度作用素 ρ も Hermite 作用素である．とくに作用素 P_j, Q_k は本質的な自己共役作用素であり（これらは閉包をとることによって自己共役となり，本質的にという制約を取り去っても一般性を失わない．よって以下自己共役作用素として扱う），Heisenberg の“**正準交換関係**”(canonical commutation relation，これを CCR と略記する）を満たしている：

$$\left.\begin{aligned} &P_jP_k-P_kP_j=0, \qquad Q_jQ_k-Q_kQ_j=0, \\ &P_jQ_k-Q_kP_j=\begin{cases} 0 & (j\neq k), \\ -iI & (j=k), \end{cases} \end{aligned}\right\} \quad (j,k=1,\cdots,m).$$

これらの式からわかるように，$j\neq k$ の場合はすべて可換となるので，作用素論的扱いが容易である．したがって，数学的に重要であるのは $j=k$ の場合，つまり各対 (P_j, Q_j) の解析であることがわかる．そこで，$m=1$ の場合，すなわち Hilbert 空間 $\mathscr{H}=L^2(\boldsymbol{R})$ における理論の展開が当面の対象となる．

議論をもとにもどすと，ひとつの状態 ρ（一般に混合）において，物理量 A を観測したとき，その観測値の平均

$$E_\rho(A)=\mathrm{tr}\,A\rho \tag{6.16}$$

を A の**期待値**という．$\rho=|x\rangle\langle x|$ $(\|x\|=1)$ のときは

$$E_x(A)\equiv E_{|x\rangle\langle x|}(A)=\langle x, Ax\rangle \tag{6.17}$$

となる．これら期待値はつねに存在するとは限らず，$A\rho$ が有界で $A\rho\in\boldsymbol{T}(\mathscr{H})$ の場合か，$x\in\mathrm{dom}(A)$ の場合に（かつ，そのときのみ）存在する．また，

$$\left.\begin{aligned} &\mathrm{var}(A;\rho)=\mathrm{tr}\,A^2\rho-(\mathrm{tr}\,A\rho)^2\geqq 0,\\ &\mathrm{var}(A;x)=\mathrm{var}(A;|x\rangle\langle x|)=\|Ax-E_x(A)x\|^2\geqq 0 \end{aligned}\right\}\tag{6.18}$$

とおき，これを A の ρ（または x）における**分散**というが，これに対して「等式 $\mathrm{var}(A;x)=0$ と，$E_x(A)$ が A の固有値で，x が A の固有ベクトルであることとは同値」であることが容易に証明される．さらに，一般に次のことがいえる．

〈定理 6.17〉 自己共役作用素 A,B が $AB=BA$ ならば，任意の $\varepsilon>0$ に対して，$x\in\mathrm{dom}(A)\cap\mathrm{dom}(B)$，$\|x\|=1$ が存在して

$$0\leqq\mathrm{var}(A;x)<\varepsilon \quad かつ \quad 0\leqq\mathrm{var}(B;x)<\varepsilon.\tag{6.19}$$

これは，A,B のスペクトル分解を用い，$\mathscr{H}$ の閉部分空間 $\mathcal{L}$ が存在し，$\mathcal{L}\subset\mathrm{dom}(A)\cap\mathrm{dom}(B)$ かつ A,B が $\mathcal{L}$ 上で有界であり，ある $y\in\mathcal{L}$ で $ABy=BAy\neq 0$ となることを使って証明できる．一方，$AB\neq AB$ ならば，上述の定理は必ずしも成立しない．つまり，任意の $\varepsilon>0$ に対して，(6.19) を満たす状態は（純粋も混合もともに）存在しないのである．

以下，状態 ρ において，$A_\rho=A-E_\rho(A)$，$B_\rho=B-E_\rho(B)$ とおく．

〈定理 6.18〉 自己共役作用素 A,B が CCR

$$AB-BA=-iI$$

を満たせば，$A_\rho\rho^{1/2}\neq 0$，$B_\rho\rho^{1/2}\neq 0$ が定義された状態 ρ に対して，

$$\mathrm{var}(A;\rho)\cdot\mathrm{var}(B;\rho)\geqq\frac{1}{4}.\tag{6.20}$$

証明 記号の簡単のため，$A_\rho=R$，$B_\rho=S$ とおく．

$$\mathrm{tr}\,RS\rho=\mathrm{tr}\,\rho RS=\mathrm{tr}(SR\rho)^*=\overline{\mathrm{tr}\,SR\rho},$$

$$2\,\mathrm{Im}(\mathrm{tr}\,RS\rho)=-i(\mathrm{tr}\,RS\rho-\mathrm{tr}\,SR\rho)=-i\,\mathrm{tr}(RS-SR)\rho=-1.$$

したがって

$$1=2|\mathrm{Im}(\mathrm{tr}\,RS\rho)|\leqq 2(\mathrm{tr}\,R^2\rho)^{1/2}\cdot(\mathrm{tr}\,S^2\rho)^{1/2},\tag{6.21}$$

すなわち

$$(\mathrm{tr}\, R^2\rho)^{1/2}\cdot(\mathrm{tr}\, S^2\rho)^{1/2} \geqq \frac{1}{2}. \qquad (6.22)$$

ここで，R は A_ρ であり，S は B_ρ であったから，

$$(\mathrm{tr}\, A_\rho{}^2\rho)^{1/2}\cdot(\mathrm{tr}\, B_\rho{}^2\rho)^{1/2} \geqq \frac{1}{2} \qquad (6.23)$$

がいえる．一方，

$$\begin{aligned}\mathrm{tr}\, A_\rho{}^2\rho &= \mathrm{tr}(A-E_\rho(A))^2\rho \\ &= \mathrm{tr}\, A^2\rho - 2(\mathrm{tr}\, A\rho)E_\rho(A) + (\mathrm{tr}\,\rho)E_\rho(A)^2 \\ &= \mathrm{tr}\, A^2\rho - (\mathrm{tr}\, A\rho)^2 \\ &= \mathrm{var}(A;\rho).\end{aligned}$$

同様に，

$$\mathrm{tr}\, B_\rho{}^2\rho = \mathrm{var}(B;\rho).$$

この2つの関係式を(6.23)に代入すると，結論の不等式(6.20)が得られる．▮

〈**定理 6.19**〉 不等式(6.20)において等号が成立する必要十分条件は，ある定数 c が存在して，

$$(A-E_\rho(A))\rho^{1/2} = c(B-E_\rho(B))\rho^{1/2}. \qquad (6.24)$$

証明 前定理の証明中の式(6.21)と(6.23)の構成を振り返ると，次のことがいえる．

"不等式(6.20)において等号が成立"

$$\Longleftrightarrow |\mathrm{Im}(\mathrm{tr}\, A_\rho B_\rho\rho)| = (\mathrm{tr}\, A_\rho{}^2\rho)^{1/2}(\mathrm{tr}\, B_\rho\rho)^{1/2}.$$

これは Schwarz の不等式における等号の成立条件に相当し，したがって

$$\Longleftrightarrow A_\rho\rho^{1/2} = cB_\rho\rho^{1/2} \Longleftrightarrow (6.24).$$ ▮

以上のことは，CCR を満たす自己共役作用素 A, B（これは当然 $AB \fallingdotseq BA$）の分散を同時に小さくすることが，〈定理 6.18〉の条件を満たすいかなる状態 ρ のもとにおいても不可能であることを示している．とくに ρ を純粋状態に限定した場合は，A, B の固有ベクトル以外の状態ベクトル $x \in \mathcal{H}, \|x\|=1$ に対して

$$\mathrm{var}(A;x)\cdot\mathrm{var}(B;x) \geqq \frac{1}{4}. \qquad (6.25)$$

つまり，2つの物理量 A, B を状態ベクトル x で観測したとき，その観測値を同時に正確に知ることは不可能なのである．

以上の2つの定理により，(6.25) の左辺を最小にする状態，すなわち等式

$$(A-E_x(A))x=c(B-E_x(B))x \tag{6.26}$$

を満たす $x\in\mathcal{H}$, $\|x\|=1$, が A, B の観測値を同時に最も良く近似する状態であることが証明されたことになる．そこで次の目標は，この x を具体的に求めることである．つまり，交換関係 CCR を満たす自己共役作用素の対 A, B を与えたとき，等式 (6.26) において $x\in\mathcal{H}$ を未知のベクトルと考えて，この方程式を解くことである．

以下，自由度1の物理系を考える．このとき，Hilbert 空間 $\mathcal{H}=L^2(\boldsymbol{R})$ 上で運動量作用素 $P=-i(\partial/\partial q)$，位置作用素 $Q=q\cdot$ を考えると，これは $\mathcal{H}$ 上の自己共役作用素の対で，CCR

$$PQ-QP=-iI$$

を満たしている．そこで，

$$(P-E_x(P))x=c(Q-E_x(Q))x \tag{6.27}$$

を満たすような $x\in\mathrm{dom}(P)\cap\mathrm{dom}(Q)$, $\|x\|=1$ の存在を示すと同時に，この状態ベクトルを具体的な波動函数として求めよう．$a=E_x(P)$, $b=E_x(Q)$ とおくと，(6.27) は微分方程式

$$\left(i\frac{\partial}{\partial q}+a\right)x=c(b-q)x \tag{6.28}$$

となり，目標の x を求めることは微分方程式 (6.28) の解を求めることである．(6.28) をかき直すと，

$$\frac{\partial x}{\partial q}=(ia+bc'-c'q)x$$

となり（ただし，$c'=-ic$），これを解くと，

$$\begin{aligned} x(q)&=c_0\exp\left(iaq+bc'q-\frac{c'q^2}{2}\right)\\ &=c_1\exp\left(-\frac{c'}{2}(q-b)^2+iaq\right). \end{aligned} \tag{6.29}$$

このようにして，実際に函数 x を求めることができたので，同時にその存在も立証されたことになる．(6.29) において，係数 c', c_1 を変形して標準型波動函数として

$$x(q)=(2\pi\sigma^2)^{-1/4}\exp\left(-\frac{(q-b)^2}{4\sigma^2}+iaq\right),\qquad \sigma=\mathrm{var}(Q;x)$$

が得られる．これが求める波動函数である．

よって，交換関係の対として P, Q を選んだとき，この物理量をひとつの状態で観測するのに最適の状態は Gauss 型の波動函数によって与えられるということがわかった．ところで，この x の Fourier 変換 $\hat{x}(p)$ を求めると，再びGauss 型函数

$$\hat{x}(p)=\left(\frac{2\sigma^2}{\pi}\right)^{1/4}\exp(-\sigma^2(p-a)^2-ib(p-a)) \qquad (6.30)$$

が得られ，この対 $(x, \hat{x})$ を **Gauss 対** (**Gaussian pair**) という．

不確定性関係を表す式 (6.20) は，信号の解析にも本質的な関係式として現れる．信号が時間軸に沿ってスカラー表示されたとき，それを観測するときの時間のゆらぎ Δt と，その信号を Fourier 変換することによって得られる周波数スペクトルのゆらぎ $\Delta\omega$ との間には，つねに関係式

$$(\Delta t)^2\cdot(\Delta\omega)^2\geq\alpha>0 \qquad (6.31)$$

が成立する．ここで，α は Fourier 変換などの単位のとり方によって定まる定数である．また，$\Delta t, \Delta\omega$ は，それぞれの信号に従属して定まる固有の正の数量であり，(6.31) は“信号のエネルギーを一定にする”という制限のもとで，信号の選び方によらずつねに成立する関係式である．すなわち，信号をある時点で観測したとき，その時間のゆらぎが Δt であれば，信号の周波数スペクトルのゆらぎを $\sqrt{\alpha}/\Delta t$ より小さくすることは不可能であり，逆に周波数スペクトルのゆらぎ $\Delta\omega$ が微小であるとき，その信号の観測による時間のゆらぎは $\sqrt{\alpha}/\Delta\omega$ より大きくなることを意味している．したがって，t と ω を同時に正確に知ることはできないのである．

不等式 (6.31) を量子系の測定に対して得られた不等式 (6.20) を使って証明しよう．以下，時間軸上で定義された，エネルギーが1である信号の全体を

$\boldsymbol{S}_1$，すなわち，

$$\boldsymbol{S}_1=\left\{x\in L^2(\boldsymbol{R});\ \|x\|^2=\int_{\boldsymbol{R}}|x(t)|^2dt=1\right\}$$

とおく．このとき各 $x\in\boldsymbol{S}_1$ に対して，x の周波数函数は x の Fourier 変換

$$\hat{x}(\omega)=\underset{T\to\infty}{\mathrm{l.i.m}}\frac{1}{\sqrt{2\pi}}\int_{-T}^{T}e^{-i\omega t}x(t)dt$$

によって表される．Fourier 解析における Plancherel の定理（第6章6.1節）によって，

$$\|x\|=\|\hat{x}\|=1$$

であるから，函数 $|x(t)|^2$ と $|\hat{x}(\omega)|^2$ は実数直線 $\boldsymbol{R}$ 上の密度函数と見なせる．これらの第一，第二モーメントを，それぞれ

$$\left.\begin{aligned}\langle t\rangle&=\int_{\boldsymbol{R}}t|x(t)|^2dt, &\langle t^2\rangle&=\int_{\boldsymbol{R}}t^2|x(t)|^2dt,\\ \langle\omega\rangle&=\int_{\boldsymbol{R}}\omega|\hat{x}(\omega)|^2d\omega, &\langle\omega^2\rangle&=\int_{\boldsymbol{R}}\omega^2|\hat{x}(\omega)|^2d\omega\end{aligned}\right\}\tag{6.32}$$

で表す．今，これらの存在を仮定し，さらに，

$$(\varDelta t)^2=\int_{\boldsymbol{R}}(t-\langle t\rangle)^2|x(t)|^2dt=\langle t^2\rangle-\langle t\rangle^2>0,$$

$$(\varDelta\omega)^2=\int_{\boldsymbol{R}}(\omega-\langle\omega\rangle)^2|\hat{x}(\omega)|^2d\omega=\langle\omega\rangle^2-\langle\omega\rangle^2>0$$

とおくと，前述の信号を観測することによるゆらぎ $\varDelta t$ や，周波数スペクトルのゆらぎ $\varDelta\omega$ は，上式で定められる．ここで，これらを運動量作用素 $P=-i(d/dt)$ と位置作用素 $Q=t\cdot$ を用いて計算する：

$$(\mathcal{F}Px)(\omega)=\omega\hat{x}(\omega),\qquad (Qx)(t)=tx(t)$$

であるから，(6.32) の各々が存在することと，$x\in\mathrm{dom}(P)\cap\mathrm{dom}(Q)$ が同等であることがわかり，

$$(\varDelta t)^2=\mathrm{var}(Q;x),\qquad (\varDelta\omega)^2=\mathrm{var}(P;x)$$

が得られる．したがって，不等式 (6.20) より

$$(\varDelta t)^2\cdot(\varDelta\omega)^2\geq\frac{1}{4}$$

が成立する．かくして，$(\Delta t)^2\cdot(\Delta\omega)^2$ が信号の不確定さの大きさを表す量と見なされ，その不確定さは，エネルギー1のいかなる信号に対しても 1/4 より小さくならないことがわかった．これは，次のように微分エントロピー形式によって表すことも可能である．

信号 $x\in \boldsymbol{S}_1$ に対して，

$$S(x)=-\int_{\boldsymbol{R}}|x(t)|^2\log|x(t)|^2dt-\int_{\boldsymbol{R}}|\hat{x}(\omega)|^2\log|\hat{x}(\omega)|^2d\omega$$

とおく．これを信号 x の時間と周波数に関する**同時エントロピー**というが，これに関して，次の定理が成り立つ．

〈**定理 6.20**〉　$x\in \boldsymbol{S}_1$，すなわち，エネルギー1の信号 x に対して，同時エントロピーはつねに，

$$S(x)\geq\log(\pi e)$$

であり，等号が成立するのは，$\{x,\hat{x}\}$ が (6.29), (6.30) によって与えられる Gauss 対の場合である．

こうして，エネルギー1の信号を観測したとき，その不確定性の大きさはつねに定数 $\log(\pi e)$ を下まわることはなく，Gauss 信号のときに最小となることがわかった．

以上，不確定性に関して種々の議論をしてきた．そのうち，信号の不確定性に関する部分はランダム性が伴わない場合であった．実は，この理論はさらに続き，ランダム信号過程（離散的および連続的）の場合に，非常に興味ある議論がなされ，不確定性関係をめぐって数学的に厳密な理論が展開されるが，ここでは触れない．

付章　数学的基礎

1.　集合と代数

この章では，集合と代数系に関する基本的結果をまとめておこう．

1.1　集　　合

ある定まった条件，または性質 P をもつ識別可能なもの $a, b, c, \cdots$ の全体を**集合**といい，$X=\{a, b, c, \cdots\}$，$X=\{a;P\}$ などと表す．この $a, b, c, \cdots$ を集合 X の**要素**，**元**，あるいは**点**といい，$a\in X$ と書き，a は X に属するという．元 a が集合 X に属さないとき $a\notin X$ とかく．なお，元を全く含まない集合を**空集合**といい，$\emptyset$ で表す．

2つの集合 X, Y において，$a\in X$ ならば $a\in Y$ となるとき，$X\subset Y$ とかき，X は Y の**部分集合**であるという．$X\subset Y$ かつ $Y\subset X$ であるとき，X と Y は全体として等しく，$X=Y$ とかく．

例 1　自然数の全体，整数の全体，有理数の全体，実数の全体，複素数の全体は，すべて集合であり，それらは $\boldsymbol{N}, \boldsymbol{Z}, \boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{C}$ で表され，$\boldsymbol{N}\subset\boldsymbol{Z}\subset\boldsymbol{Q}\subset\boldsymbol{R}\subset\boldsymbol{C}$ の関係を有する．

$X\cup Y\equiv\{a;a\in X$ または $a\in Y\}$を X と Y の**和集合**，$X\cap Y\equiv\{a;a\in X$ かつ $a\in Y\}$ を X と Y の**共通集合**という．$X\cap Y=\emptyset$ のとき，X と Y は**たがいに素**であるといい，このとき $X\cup Y=X+Y$ とかき，これを X と Y の**直和**という．

集合 X の部分集合 Y の X に関する補集合 Y^c とは**差集合** $X\backslash Y\equiv\{x\in X;x\notin Y\}$ のことである（これは $X-Y$ と表されることもある）．X の2つの部分集合 A, B の**対称差** $A\triangle B$ とは $(A\backslash B)\cup(B\backslash A)$ のことである．

ある与えられた集合 $J(\neq\emptyset)$ の各元 $\alpha\in J$ に集合 X_α が対応しているとき，この集合の全体 $\{X_\alpha;\alpha\in J\}$ において，和集合および共通集合を各々 $\bigcup_{\alpha\in J}X_\alpha$，$\bigcap_{\alpha\in J}X_\alpha$ で表す．

$A, B, C, A_\alpha(\alpha\in J)$ 等を集合 X の部分集合とする．このとき，次の関係が成立する．

（1）　**交換律**：$A\cup B=B\cup A$，$A\cap B=B\cap A$.

（2）　**結合律**：$(A\cup B)\cup C=A\cup(B\cup C)$，$(A\cap B)\cap C=A\cap(B\cap C)$.

（3）　**分配律**：$(A\cup B)\cap C=(A\cap C)\cup(B\cap C)$，$(A\cap B)\cup C=(A\cup C)\cap(B\cup C)$.

（4）　**双補律**：$A\cap A^c=\emptyset$，$A\cup A^c=X$，$(A^c)^c=A$，$X^c=\emptyset$.

（5）　**de Morgan の法則**：$(\bigcup_{\alpha\in J}A_\alpha)^c=\bigcap_{\alpha\in J}A_\alpha{}^c$，$(\bigcap_{\alpha\in J}A_\alpha)^c=\bigcup_{\alpha\in J}A_\alpha{}^c$.

$\{A_n;n\in\boldsymbol{N}\}$ を X の部分集合の列とするとき，$\overline{\lim_{n\to\infty}}A_n\equiv\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k$，$\underline{\lim_{n\to\infty}}A_n\equiv\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}A_k$ を各々 $\{A_n\}$ の**上極限集合**および**下極限集合**という．さらに，$\overline{\lim}A_n=\underline{\lim}A_n$ であ

るとき，これを単に $\lim A_n$ とかき，$\{A_n\}$ の**極限集合**という．なお，$\overline{\lim}=\limsup$，$\underline{\lim}=\liminf$ とかくこともある．これらの極限に関して

1° $\underline{\lim}\,A_n\subset\overline{\lim}\,A_n$，$(\underline{\lim}\,A_n{}^c)^c=\overline{\lim}\,A_n$，$(\overline{\lim}\,A_n{}^c)^c=\underline{\lim}\,A_n$.

2° $A_1\subset A_2\subset\cdots$（$A_n\uparrow$ とかく）ならば，$\lim A_n=\cup A_n$．また，$A_1\supset A_2\supset\cdots$（$A_n\downarrow$ とかく）ならば $\lim A_n=\cap A_n$.

1.2 写　　像

集合 X の各元に集合 Y のひとつの元を対応させる規則 φ を X から Y への**写像**または**函数**といい，$\varphi:X\to Y$ で表す．このとき X を φ の**定義域**（φ が X 全体で定義されていず，X のある部分集合上で定義されているとき，その定義域を $\mathrm{dom}\varphi$ で表す），$\varphi(X)\equiv\{\varphi(x);x\in X\}$ を φ の**値域**という（$\varphi(X)$ を $\mathrm{ran}\varphi$ とかく）．$A\subset X$，$B\subset Y$ に対して，$\varphi(A)\equiv\{\varphi(x);x\in A\}$，$\varphi^{-1}(B)\equiv\{x\in X;\varphi(x)\in B\}$ を各々 A, B の φ による**像**および**原像**とよぶ．なお，任意の $y\in\varphi(X)$ に対して，一意に $x\in X$ が決まり，$y=\varphi(x)$ であるとき，φ は**単射**(injection) あるいは**1対1写像**であるといい，このとき写像 φ^{-1}: $\varphi(X)\to X$ が存在して，$\varphi^{-1}(\varphi(x))=x$ $(\forall x\in X)$，$\varphi(\varphi^{-1}(y))=y$ $(\forall y\in\varphi(X))$ を満たす．この φ^{-1} を φ の**逆写像**という．写像 $\varphi:X\to Y$ が $Y=\varphi(X)$ を満たすとき，φ は X から Y への**全射**(surjection) あるいは**上への写像**であるといわれる．また，φ が全射でかつ単射であるとき**全単射**（bijection）という．

集合 X, Y において，$X\times Y\equiv\{(x,y);x\in X, y\in Y\}$ を X と Y の**直積集合**，元 (x,y) を**順序対**とよぶ．また，集合族 $\{X_\alpha;\alpha\in J\}$ に対して，各 $\alpha\in J$ に $x_\alpha\in X_\alpha$ を対応させる写像 φ(i.e., $\varphi(\alpha)=x_\alpha$) の全体を $\prod_{\alpha\in J}X_\alpha$ とかき，$\{X_\alpha\}$ の**直積**という．この φ は $(x_\alpha)_{\alpha\in J}$ あるいは $\prod_{\alpha\in J}x_\alpha$ で表される（同一視できる）から，x_α を φ の **α-座標**とよぶ．さらに，対応 $P_\alpha:\prod x_\alpha\to x_\alpha$ を $\prod X_\alpha$ から X_α への**射影**という．とくに，$X_\alpha=X$ $(\forall\alpha\in J)$ のとき，$\prod X_\alpha$ を X^J とかき，X を底とし，J を指標とする**ベキ（配置）集合**とよぶ．これは J から X への写像の全体である．

問 1　$X=\{1,2,3,4\}$ のとき，次の各順序対の集合は X から X への写像となるか?：（1）$\varphi_1\equiv\{(2,3),(1,4),(2,1),(3,2),(4,4)\}$，（2）$\varphi_2\equiv\{(3,1),(4,2),(1,1)\}$，（3）$\varphi_3\equiv\{(2,1),(3,4),(1,4),(4,4)\}$，ここで順序対 (a,b) は対応 $a\to b$ を表す．

問 2　$A_1, A_2\subset X$，$B_1, B_2\subset Y$，$\varphi:X\to Y$ に対して，次の事実を証明せよ：（1）$\varphi(A_1\cup A_2)=\varphi(A_1)\cup\varphi(A_2)$，$\varphi(A_1\cap A_2)\subset\varphi(A_1)\cap\varphi(A_2)$；（2）$A_1\supset A_2$ のとき，$\varphi(A_1\backslash A_2)\supset\varphi(A_1)\backslash\varphi(A_2)$；（3）$\varphi^{-1}(B_1\cup B_2)=\varphi^{-1}(B_1)\cup\varphi^{-1}(B_2)$，$\varphi^{-1}(B_1\cap B_2)=\varphi^{-1}(B_1)\cap\varphi^{-1}(B_2)$；（4）$B_2\subset B_1$ のとき $\varphi^{-1}(B_1\backslash B_2)=\varphi^{-1}(B_1)\backslash\varphi^{-1}(B_2)$；（5）$A_1\subset\varphi^{-1}(\varphi(A_1))$，$\varphi(\varphi^{-1}(B_1))=B_1\cap\varphi(X)$，$\varphi(A_1\cap\varphi^{-1}(B_1))=\varphi(A_1)\cap B_1$.

1.3 同値関係

集合 X と Y の**関係**とは部分集合 $R\subset X\times Y$ のことであり，$x\in X$ と $y\in Y$ が **R-関**

係にあるとは，$(x, y)\in R$ となる場合をいう．$(x, y)\in R$ は xRy，$x=y(\mathrm{mod}\, R)$，$x\overset{R}{\sim}y$ などと表すこともある．

関係 $R\subset X\times Y$ が X における**同値関係**であるとは，次の条件が成立するときをいう．

（1） **反射的**：任意の $x\in X$ に対して $(x, x)\in R$.

（2） **対称的**：$(x, y)\in R$ ならば $(y, x)\in R$.

（3） **推移的**：$(x, y), (y, z)\in R$ ならば $(x, z)\in R$.

ところで，$[x]\equiv\{y\in X; (x, y)\in R\}$ を $x\in X$ の**同値類**とよぶが，X はたがいに素な同値類の和集合として表すことができる．この同値類の全体を X/R とかき，**商集合**という．なお，X のある同値関係を "R" の代わりに "$\sim$" で表すこともあるが，その場合，$x\sim y$，$X/\sim$ などの記法が用いられる．

問 3 $X=\{2, 3, 4, 5\}$ と $Y=\{3, 6, 7, 10\}$ の関係を "$x\in X$ が $y\in Y$ を割り切る" によって定める．この関係を表す順序対の集合 R をかき下せ．

問 4 $R^{-1}\equiv\{(y, x); (x, y)\in R\}$ を**逆関係**というが，R が推移的なとき，R^{-1} も推移的であることを示せ．

問 5 集合 $\boldsymbol{Z}$ において，"$xRy\overset{\mathrm{def}}{\Longleftrightarrow}x-y=0(\mathrm{mod}\, 3)$" で関係 R を定める．このとき，R が同値関係であることを示し，同値類の集合 $\boldsymbol{Z}/R$ を求めよ．

1.4 濃　　度

2つの集合 X, Y の間に全単射が存在するとき，X と Y とは**対等**である（$X\sim Y$ とかく），または，同じ**濃度**あるいは**基数**（これを $|X|$ で表す）をもつという．

$X\sim\{1, 2, 3, \cdots, n\}$ ならば $|X|=n$ で，さらに $n<\infty$ のとき X は**有限集合**であるといわれる．また，$X\sim Y$ となる真の部分集合 Y（i. e., $Y\subsetneqq X$）が存在するとき，X は**無限集合**であるという．とくに，$X\sim\boldsymbol{N}$ であるとき，X は**可算**（または**可付番**）であるといい，このとき，X の濃度を $\mathfrak{a}$ または $\aleph_0$（アレフゼロと読む）で表す．$X\sim[0, 1]$ のとき，X は連続体の濃度をもつといい，$|X|=\mathfrak{c}$ または $\aleph$ と記す．さらに，区間 $[0, 1]$ 上の実数値函数の全体の集合の濃度を $\mathfrak{f}$ で表す．

1° 上記の "$\sim$" は同値関係を与える．

2° $A_n\cap A_n=\emptyset$ $(n\neq m)$，$A_n\sim\boldsymbol{N}\Longrightarrow\bigcup_{n\in\boldsymbol{N}}A_n\sim\boldsymbol{N}$.

3° $\boldsymbol{N}\times\boldsymbol{N}\sim\boldsymbol{N}$，$\boldsymbol{Q}\sim\boldsymbol{N}$.

X が Y の部分集合と対等であるとき，$|X|\leqq|Y|$ とかき，$|X|$ は $|Y|$ より小さいと読む．また，$|X+Y|\equiv|X|+|Y|$，$|X\times Y|\equiv|X||Y|$，$|X^Y|\equiv|X|^{|Y|}$ と定めると，次の各性質が容易に証明される．

4° 任意の濃度 α, β, γ に対して次が成立：（1）$\alpha+\beta=\beta+\alpha$，（2）$\alpha\beta=\beta\alpha$，（3）$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$，（4）$(\alpha\beta)\gamma=\alpha(\beta\gamma)$，（5）$\alpha(\beta+\gamma)=\alpha\beta+\alpha\gamma$，（6）$\alpha^\beta\alpha^\gamma=\alpha^{\beta+\gamma}$，（7）$(\alpha^\beta)^\gamma=\alpha^{\beta\gamma}$，（8）$(\alpha\beta)^\gamma=\alpha^\gamma\beta^\gamma$.

5° $\aleph_0=n+\aleph_0=\aleph_0+\aleph_0=n\aleph_0=\aleph_0{}^2=\cdots=\aleph_0{}^n$ $(\forall n\in\boldsymbol{N})$.

6° **Cantor の定理**：$|X|<2^{|X|}$.

7° **Schröder-Bernstein の定理**：$|X|\leqq|Y|$かつ$|Y|\leqq|X|$ であれば $|X|=|Y|$.

8° $2^{\aleph_0}=n^{\aleph_0}=\cdots=\aleph_0{}^{\aleph_0}=\aleph$, $2^{\aleph}=n^{\aleph}=\aleph_0{}^{\aleph}=\aleph^{\aleph}=\mathfrak{f}$.

9° **比較定理**：任意の濃度 α,β に対して，$\alpha<\beta$, $\alpha=\beta$, $\alpha>\beta$ のいずれかが成立.

ここで，上記の Cantor の定理と 8° から，$\aleph_0<2^{\aleph_0}=\aleph$ はわかるが，では，$\aleph_0$ と $\aleph$ の間の濃度をもつ集合は存在するかという疑問が起こる．この問に関して，最初は，そのような集合は存在しない（**連続体仮説**）と考えられていたが，Cohenは この仮説が集合論の公理系と独立であることを証明した．すなわち，$\aleph_0$ と $\aleph$ の間の濃度をもつ集合が存在してもしなくても，集合論の公理系とは何ら矛盾しないのである．

1.5 順　　序

関係 $R\subset X\times X$ が**順序**（半順序ということもある）であるとは

（1） **反射的**：任意の $x\in X$ に対して $(x,x)\in R$;

（2） **反対称的**：$(x,y)\in R$ かつ $(y,x)\in R$ のとき $x=y$;

（3） **推移的**：(x,y), $(y,z)\in R$ のとき $(x,z)\in R$;

を満たす場合をいう．以下，順序関係 R を“$\prec$”で表す．すなわち，$x\prec y\Longleftrightarrow(x,y)\in R$. 順序 $\prec$ をもつ集合 X を組 $(X,\prec)$ で表すことが多い．

$(X,\prec)$ において，$x\in X$ が任意の $y\in X$ に対して，$y\prec x$ ($x\prec y$) となるとき，x を X の**最大（小）元**という．また，$x\in X$ が**極大（小）元**であるとは $x\prec y$ ($y\prec x$), $y\in X$, ならば $x=y$ となる場合である．さらに，$A\subset X$ と $x\in X$ に対して，すべての元 $a\in A$ が $a\prec x$ ($x\prec a$) を満たすとき，x を A の**上（下）界**という．A の上（下）界の中に最小（大）元が存在すれば，それを上（下）限といい，$\sup A$ ($\inf A$) あるいは l. u. b. A (g. l. b. A) で表す．

$(X,\prec)$ において，任意の $x,y\in X$ に対して，$z\in X$ が存在して $x\prec z$, $y\prec z$ となるとき，$(X,\prec)$ を**有向集合**という．また，任意の $x,y\in X$ に対して，$x\prec y$ または $y\prec x$ のどちらかが必ず成立するとき，$(X,\prec)$ は**全順序集合**であるという．全順序集合では極大元＝最大元.

全順序集合 $(X,\prec)$ において，X の任意の部分集合 $Y(\neq\emptyset)$ が最小元をもつとき，$(X,\prec)$ は**整列集合**であるという（たとえば，集合 $\boldsymbol{N}$）．順序集合 X の任意の全順序部分集合がすべて上に有界であるとき，X を**帰納的順序集合**という．

さて，次の公理は現代解析において非常に重要な役割を果たしているものである．

〈選択公理〉 空でない集合の集合族 $\{X_\alpha;\alpha\in J\}$ において，各集合 X_α から同時にひとつずつ元を取り出すことができる．

この公理は，以下の3つの命題と同値であることが知られている．

〈Zermelo の公準〉 たがいに素な空でない集合の集合族 $\{X_\alpha;\alpha\in J\}$ において，$\bigcup_{\alpha\in J}X_\alpha$ の部分集合 Y で $X_\alpha\cap Y$ $(\forall\alpha\in J)$ がただひとつの元からなるようなものが存在する．

〈Zorn の補題〉 空でない帰納的順序集合は少なくともひとつ極大元をもつ．

〈整列可能定理〉 任意の集合は適当な順序によって整列化できる．

問 6 $X=\{a,b,c,d,e,f\}$ に図1のように順序を与える．このとき，次の問に答えよ．

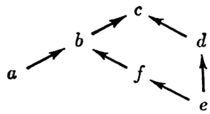


図 1

（1） X の極大元，極小元，最大元，最小元を求めよ．

（2） $\mathcal{A}$ を X の全順序部分集合の族とし，$\mathcal{A}$ に包含関係で順序を与える．このとき $\mathcal{A}$ の極大元，極小元，最大元，最小元を求めよ．

問 7 2つの順序集合 X, Y において，X から Y への全単射 φ が存在して，$x_1\prec x_2 \Rightarrow \varphi(x_1)\prec\varphi(x_2)$ となるとき，X と Y は**相似**であるといい，$X\backsim Y$ で表すと，この"$\backsim$"は同値関係になることを示せ．また，X が全順序集合で，$X\backsim Y$ であれば，Y も全順序集合であることを証明せよ．

問 8 線形空間 X が基底をもつことを Zorn の補題を使って証明せよ．

問 9 $\boldsymbol{R}^2$ に順序を"$(x,y)\prec(x',y')\Longleftrightarrow x<x'$ または $x=x'$ のときは $y\leqq y'$"で入ると，$\boldsymbol{R}^2$ は順序集合であるが全順序集合ではないことを示せ．

1.6 群

演算 $\circ$ が定義されている集合 G が**半群** (semigroup) であるとは，G の任意の元 a,b,c に対して，$(a\circ b)\circ c=a\circ(b\circ c)$（**結合律**）が満たされるときをいう．半群 G の元 e が $a\circ e=e\circ a=a$ を満たすとき，e を G の**単位元**という．単位元 e をもつ半群 G が，その任意の元 a に対して，$a\circ b=b\circ a=e$ となる元 $b\in G$ をもつとき，G は演算 $\circ$ に関して**群** (group) をなすといい，この b を a^{-1} で表し，a の**逆元**とよぶ．とくに，群 G の任意の元 a,b に対して $a\circ b=b\circ a$ であるとき，G は**可換群**（または **Abel 群**）であるという．

群 G において，演算 $\circ$ が $a\circ b=a+b$ で定められているとき，G は**加（法）群**であるといい，$a\circ b=ab$ のとき**乗法群**であるという．

G を2個以上の元からなる集合 X からそれ自身への全単射の全体とし，$\circ$ を写像の合成 (i.e., $f\circ g(x)\equiv f(g(x))$) とすると，$G$ は非可換群となり，この群を**置換群**という．とくに，$X=\{1,2,\cdots,n\}$ であるとき，G は n 次**対称群**とよばれる．さらに，$|G|=n<\infty$ のとき，G を**有限群**といい，この n を G の**位数**という．

群 G の部分集合 H が**部分群**であるとは，任意の $a,b\in H$ に対して，$a\circ b\in H$ かつ $a^{-1}, b^{-1}\in H$ の場合をいう．

群 G の部分群 H と $g\in G$ に対して，集合 $g\circ H\equiv\{g\circ h;h\in H\}$ を H の**左剰余類** (left coset) といい，$H\circ g\equiv\{h\circ g;h\in H\}$ を H の**右剰余類** (right coset) という．"$a\sim b\overset{\text{def}}{\Longleftrightarrow}$

$a\circ b^{-1}$(resp. $a^{-1}\circ b)\in H$” で関係～を定めると，これは同値関係になり，その同値類は $[a]\equiv H\circ a$(resp. $a\circ H$) となる．すべての元 $g\in G$ に対して，$g\circ H=H\circ g$ であるとき，H を G の**正規部分群**という．正規部分群 H のすべての剰余類の集まりを G の**商群**といい，G/H で表す．なお，G/H の演算☆は $(a\circ H)\star(b\circ H)=(a\circ b)\circ H$ である．

問 1　n 次を正則行列の全体 $GL(n, C)$ は乗法群であることを証明せよ．

問 2　H が群 G の部分群であるとき次が成立することを証明せよ．（1）$H\circ a=H\circ b \Longleftrightarrow a^{-1}\circ H=b^{-1}\circ H$ $(\forall a, b\in G)$．（2）$H\circ a$ を $a^{-1}\circ H$ に写す写像は，H の左剰余類全体と右剰余類全体との間に1対1の対応を与える（それらの剰余類の個数を r とすると，$r=(G:H)$ とかき，H の G における**指数**という）．

1.7 環と体

集合 R に2つの演算，加法 $a+b$ と積 ab $(\forall a, b\in R)$ が定義されていて，（1）R は “+” のもとで群をなし，（2）任意の $a, b, c\in R$ に対して，$ab\in R$，$(ab)c=a(bc)$，$a(b+c)=ab+bc$，$(a+b)c=ac+bc$，を満たすとき，R は**環** (ring) であるという．なお，加法に関する単位元は0とかき，$a\in R$ の逆元は $-a$ とかく．さらに，積に関して可換である環を**可換環**という．積に関する単位元 e をもつ環に対して，ある元 a が逆元 a^{-1} をもてば，この元 a は**正則**あるいは**可逆**であるとよばれる．単位元をもつ環 R がその0でないすべての元に対して逆元をもつとき，環 R は**体** (field) であるという．

環 R の部分環 $\mathcal{I}$ が $R\mathcal{I}\equiv\{Ra; a\in\mathcal{I}\}\subset\mathcal{I}$ であるとき，$\mathcal{I}$ を R の**左イデアル** (left ideal) であるといい，$\mathcal{I}R\subset\mathcal{I}$ のとき**右イデアル**であるという．さらに，左かつ右イデアルである $\mathcal{I}$ を**両側イデアル**という．とくに，$\mathcal{I}\subsetneqq R$ であるイデアル $\mathcal{I}$ を**真のイデアル**とよぶ．真の両側イデアル $\mathcal{I}$ において，同値関係～を “$a\sim b \overset{\text{def}}{\Longleftrightarrow} a-b\in\mathcal{I}$” で導入すると，$R/\mathcal{I}\equiv\{[a]=a+\mathcal{I}; a\in R\}$ は演算 $[a][b]=[ab]$，$[a]+[b]=[a+b]$ に関して環をなす．この環を**商環**とよぶ．

1.8 同型写像

$(X_1, \circ), (X_2, \star)$ を2つの群（環；以下かっこの中は環の場合）として，写像 $\varphi: X_1\rightarrow X_2$ が $\varphi(a\circ b)=\varphi(a)\star\varphi(b), \forall a, b\in X_1$ ($\varphi(ab)=\varphi(a)\varphi(b)$ かつ $\varphi(a+b)=\varphi(a)+\varphi(b)$) を満たすとき，$\varphi$ を**準同型写像** (homomorphism) とよび，X_1 と X_2 は**準同型**であるといい，$X_1\simeq X_2$ とかく．とくに，φ が全単射であるとき，φ を**同型写像** (isomorphism) とよび，X_1 と X_2 は**同型**であるといい，$X_1\cong X_2$ とかく．

〈同型定理〉　φ: $X_1\rightarrow X_2$ を準同型写像とする．このとき，次が成立する．

（1）　φ の核 $\ker\varphi\equiv\{a\in X_1; \varphi(a)=e_2(\varphi(a)=0_2)\}$ は正規部分群（両側イデアル）．

（2）　$\varphi(X_1)$ は X_2 の部分群（環）である．（3）　$X_1/\ker\varphi\cong\varphi(X_1)$．

（4）　逆に，K_1 が X_1 の正規部分群（両側イデアル）であるとき，$\psi(a)=a\circ K_1$

$(\psi(a)=a+K_1)$ とすると，ψ は X_1 から X_1/K_1 への準同型写像である．

（5） K_2 を X_2 の正規部分群（両側イデアル）とし，$K_1=\varphi^{-1}(K_2)$ とおくと，K_1 および $K_1/\ker\varphi$ はそれぞれ X_1 および $X_1/\ker\varphi$ の正規部分群（両側イデアル）であり，

$$(X_1/\ker\varphi)/(K_1/\ker\varphi)\cong X_1/K_1\cong\varphi(X_1)/K_2.$$

問 3 n 次の対称群 S_n の位数を求めよ．また，S_n から $\{\pm1\}$ への写像 σ は準同型であることを示せ（なお，$A_n=\ker\sigma$ を**交代群**とよぶ）．

問 4 体 K を係数にもつ x の多項式の全体 $K[x]$ は**整域**（すなわち，零因子をもたない可換環）であることを示せ．

問 5 可換環 R において，$\mathcal{I}\subsetneqq R$ が**極大イデアル**（すなわち，$\mathcal{I}\subset\mathcal{I}_1\subsetneqq R$, $\mathcal{I}_1$ がイデアルならば $\mathcal{I}_1=\mathcal{I}$）であることと，$R/\mathcal{I}$ が体であることは同値であることを証明せよ．

1.9 束

順序集合 $(L,\prec)$ の元 $x,y\in L$ に対して，$\{x,y\}$ の上限および下限が存在するとき，それらを $x\vee y\equiv\sup\{x,y\}$，$x\wedge y\equiv\inf\{x,y\}$ とかき，それぞれ x と y の**結び**（join）および**交わり**（meet）と**よぶ**．さらに，L の任意の 2 つの元が $\vee$ と $\wedge$ をもつとき，L は**束**（lattice）であるという．束 L はその任意の x,y,z に対して次の性質を有する．

（1） **結合律**：$(x\vee y)\vee z=x\vee(y\vee z)$，$(x\wedge y)\wedge z=x\wedge(y\wedge z)$.

（2） **交換律**：$x\vee y=y\vee z$，$x\wedge y=y\wedge z$.

（3） **べき等律**：$x\vee x=x$，$x\wedge x=x$.

（4） **吸収律**：$x\vee(y\wedge x)=(x\vee y)\wedge x=x$.

ところで，束 L が

（5） **モジュラー律**：$x\wedge y=y\Rightarrow x\wedge(y\vee z)=y\vee(x\wedge z)$ を満たすとき，L を**モジュラー束**という．また（5）より強い条件

（6） **分配律**：$x\wedge(y\vee z)=(x\wedge y)\vee(x\wedge z)$，$x\vee(y\wedge z)=(x\vee y)\wedge(x\vee z)$ を満たすとき，L を**分配束**という．

束 L の任意の元 x に対して，$x\wedge0=0$，$x\wedge1=x$ となる元 0 を**零元**，1 を**単位元**とよぶが，これらの元は存在すれば一意である．$0,1\in L$ のとき，$x\wedge x^c=0$，$x\vee x^c=1$ $(\forall x\in L)$ となる x^c を x の**補元といい**，すべての元が補元をもつ束を**直補束**，直補束でかつ分配束である束を **Boole 束**または **Boole 代数**という．

Boole 束 $(L,\prec)$ の任意の元 $x,y\in L$ に対して，次が成立する：（1） $x\vee0=x$，$x\wedge1=x$；（2） $x\vee1=1$，$x\wedge0=0$；（3） $(x^c)^c=x$；（4） $x\vee y=1$ かつ $x\wedge y=0$ ならば $x=y^c$；（5） $(x\vee y)^c=x^c\wedge y^c$；（6） $(x\wedge y)^c=x^c\vee y^c$.

2 つの有限 Boole 束 L_1 と L_2 は，その元の個数が等しければ同型である（すなわち，全単射 $\varphi:L_1\to L_2$ が存在して，$\varphi(x\vee y)=\varphi(x)\vee\varphi(y)$，$\varphi(x\wedge y)=\varphi(x)\wedge\varphi(y)$，$\varphi(x^c)=\varphi(x)^c$）．したがって，任意の有限 Boole 束 L に対して，有限集合 X が存在して，X のべき集合 2^X（X の部分集合の全体）と L は同型となる．

2. 位相空間

ある与えられた集合において，要素間や集合間に近さの概念が付与されているとき，その集合を**位相空間**といい，またその集合は位相構造をもつという．これは非常に重要な概念で，本書でも頻繁に登場する．ここでは，要点をしぼり説明する．この章では考える集合を $X, Y, A, B, \cdots$ などで表す．

2.1 近傍系

与えられた集合 X において，各 $x \in X$ に対して X の部分集合族 $\mathfrak{N}(x)$ が対応していて，それが次の条件を満たす：

(**N.1**) 各 $U(x) \in \mathfrak{N}(x)$ に対して $x \in U(x)$.

(**N.2**) 各 $U(x) \in \mathfrak{N}(x)$，各 $y \in U(x)$ に対して $V(y) \in \mathfrak{N}(y)$ が存在して

$$V(y) \subset U(x).$$

(**N.3**) 各 $U(x)$，$V(x) \in \mathfrak{N}(x)$ に対して $W(x) \in \mathfrak{N}(x)$ が存在して

$$W(x) \subset U(x) \cap V(x).$$

この $\mathfrak{N}(x)$ を x の**近傍系**，各 $U(x) \in \mathfrak{N}(x)$ を x の**近傍**という．

2.2 開集合族，位相

各元 x に対して近傍系 $\mathfrak{N}(x)$ が付与されている X において，部分集合 $G \subset X$ が**開集合**であるとは，$G = \emptyset$ または $G \neq \emptyset$ で，各 $x \in G$ に対して $U(x) \in \mathfrak{N}(x)$ が存在して $U(x) \subset G$ のときをいう．開集合の全体からなる族を $\mathcal{O}$ で表す．このとき明らかに各 $x \in X$ に対して $\mathfrak{N}(x) \subset \mathcal{O}$.

1° 開集合族 $\mathcal{O}$ は次の条件を満足する：

(**O.1**) $\emptyset \in \mathcal{O}$，$X \in \mathcal{O}$,

(**O.2**) $G_1, G_2 \in \mathcal{O} \Rightarrow G_1 \cap G_2 \in \mathcal{O}$, （乗法性）

(**O.3**) $\{G_j ; j \in J\} \subset \mathcal{O} \Rightarrow \bigcup_{j \in J} G_j \in \mathcal{O}$. （完全加法性）

ここで，J は任意の濃度をもった指標 (index) 集合である．

2° 与えられた集合 $X (\neq \emptyset)$ の部分集合族 $\mathcal{O}$ が条件 (O.1～3) を満たすとき，各 $x \in X$ に対し $\mathfrak{N}_1(x) = \{G \in \mathcal{O} ; x \in G\}$ とおくと，これは条件 (N.1～3) を満たす．

3° 2° で $\mathcal{O}$ が定めた $\mathfrak{N}_1(x)$ に対して開集合 G の全体を $\mathcal{O}_1$ とおくと $\mathcal{O} = \mathcal{O}_1$.

各元 x に近傍系 $\mathfrak{N}(x)$ が付与された X（これを $\mathfrak{N}(X)$ で表す）において，一意に定まる開集合族 $\mathcal{O}$ を X の**位相**という．近傍系も付記していう場合は，**近傍による位相**，または**近傍系 $\mathfrak{N}(X)$ に対応する（またはによる）位相**という．

4° X に位相を定めるのに，はじめから条件 (O.1～3) を満たす族 $\mathcal{O}$ を与え，近傍系に関わりなくこの $\mathcal{O}$ を位相と名付けてもよい．これは，2° によって近傍系 $\mathfrak{N}_1(x)$ を定めれば，3° によって同じ $\mathcal{O}$ を得ることによる．したがって，また近傍系自身を位相と称してもよい．近傍系や開集合族を**位相構造**ともいう．

以下，X を単に**位相空間といったとき，同時に** $\mathfrak{N}(X)$ **と** $\mathcal{O}$ **を考える**場合が多い．近傍による位相 $\mathfrak{N}(X)$ が与えられているとき，もうひとつの近傍系 $\mathcal{B}(x)$ が各 $x\in X$ で $\mathcal{B}(x)\subset\mathfrak{N}(x)$ かつ次の条件を満たすとする：任意の $U(x)\in\mathfrak{N}(x)$ に対して $V(x)\in\mathcal{B}(x)$ が存在して $V(x)\subset U(x)$．このとき $\mathcal{B}(x)$ を**近傍基**という．

位相 $\mathcal{O}$ が与えられているとき，$\mathcal{O}$ の部分族 $\mathcal{B}$ が次の条件を満たすとする：$x\in X$ と $G\in\mathcal{O}$ が $x\in G$ ならば，$B\in\mathcal{B}$ が存在して $x\in B\subset G$，このとき $\mathcal{B}$ を**位相基**または単に**基**という．

5° 近傍系と，近傍基は同一の位相を定義する．

6° 位相 $\mathcal{O}$ の部分族 $\mathcal{B}\subset\mathcal{O}$ に対して次の条件は同等である：（1）$\mathcal{B}$ は位相基である．（2）$\mathcal{O}$ の任意の元は $\mathcal{B}$ の元の和として表される．

7° 近傍系 $\mathfrak{N}(x)$，$x\in X$ や位相 $\mathcal{O}$ は，それ自身が近傍基であり位相基である．

各元 $x\in X$ が高々可算個の要素からなる近傍基をもつとき，X は**第一可算公理**を満たすといい，位相 $\mathcal{O}$ が高々可算個の基をもつとき X は**第二可算公理**を満たすという．第二可算公理⇒第一可算公理である．

位相空間の簡単な例：（1） $X=\boldsymbol{R}$（実数直線）．各 $x\in X$ の ε-近傍を $U_\varepsilon(x)=\{y\in R;\ |x-y|<\varepsilon\}$ とすれば $\mathfrak{N}(x)\equiv\{U_\varepsilon(x);\varepsilon>0\}$ は近傍系となり，したがってこれは位相空間となる．ここで $B_n(x)=U_{1/n}(x)$ とおくと $\mathcal{B}(x)\equiv\{B_n(x);\ n=1,2,\cdots\}$ は可算近傍基である．また $\mathcal{B}=\{B_n(r);r\in\boldsymbol{Q}$（有理数），$n=1,2,\cdots\}$ は位相基である．（2） $X=\boldsymbol{R}^n$（n-次元 Euclid 空間)．（1）と同様に $U_\varepsilon(x)$, $B_n(x)$, $\mathcal{B}(x)$, $\mathcal{B}$ 等を定められる．

2.3 位相の比較，強・弱

X の2つの位相 $\mathcal{O}_1,\mathcal{O}_2$ に対して，集合族として $\mathcal{O}_1\subset\mathcal{O}_2$ のとき，$\mathcal{O}_1$ は $\mathcal{O}_2$ より**粗または弱い**，あるいは $\mathcal{O}_2$ は $\mathcal{O}_1$ より**密または強い**位相であるという．当然近傍系による場合も同様で，近傍系 $\mathfrak{N}_1(x),\mathfrak{N}_2(x)(x\in X)$ において，各 $x\in X$ と $U(x)\in\mathfrak{N}_1(x)$ に対して，$V(x)\in\mathfrak{N}_2(x)$ が存在して $V(x)\subset U(x)$ が成立することと，$\mathfrak{N}_1(x)$ による位相より $\mathfrak{N}_2(x)$ による位相の方が強いことは同等である．同時に弱でありかつ強であるとき，同じ位相が定義されるという．

各 $x\in X$ に対し $\mathfrak{N}_0(x)=\{X\}$，$\mathfrak{N}_d(x)=\{A\subset X;x\in A\}$ とおき，各々に対応する位相を $\mathcal{O}_0,\mathcal{O}_d$ とすると，$\mathcal{O}_0=\{\emptyset,X\}$，$\mathcal{O}_d=\{X$ のすべての部分集合からなる族$\}$．このとき $\mathcal{O}_0$ を**極粗位相**，$\mathcal{O}_d$ を**離散位相**という．極粗位相は最も弱い位相であり，離散位相は最も強い位相である．

2.4 閉集合，集積点，閉包

位相空間 $(X,\mathcal{O})$ において，部分集合 $F\subset X$ の補集合が $F^c\equiv X\backslash F\in\mathcal{O}$ であるとき，F を**閉集合**という．X の閉集合の全体の族を $\mathcal{F}$ で表す．

1° この閉集合族 $\mathcal{F}$ は次の条件を満たす：

(**F.** 1)　$\emptyset\in\mathcal{F}$, $X\in\mathcal{F}$,

(**F.** 2)　$F_1, F_2\in\mathcal{F}\Longrightarrow F_1\cup F_2\in\mathcal{F}$,　(加法性)

(**F.** 3)　$\{F_j; j\in J\}\subset\mathcal{F}\Longrightarrow \bigcap_{j\in J}F_j\in\mathcal{F}$.　(完全乗法性)

逆に，この (F. 1〜3) を満たす $\mathcal{F}$ が与えられているとき，$\mathcal{O}=\{F^c; F\in\mathcal{F}\}$ は条件 (O. 1〜3) を満たし，X に位相が定義される.

以下，位相空間 X において，部分集合 $A, B\subset X$，元 $x\in X$ などを考える.

x が A の**集積点**$\overset{\text{def}}{\Longleftrightarrow}$任意の $U(x)\in\mathfrak{N}(x)$ に対して，$(U(x)\setminus\{x\})\cap A\neq\emptyset$. A の集積点の全体を A^d とかき，A の**導集合**という.

A を含む最小の閉集合を $\bar{A}$ あるいは A^a で表し，それを A の**閉包**という．すなわち，$\bar{A}=\cap\{F\in\mathcal{F}; A\subset F\}$. $\bar{A}$ の元を A の**触点**という．A に含まれる最大の開集合を A° で表し，それを A の**開核**という，すなわち，$A^\circ=\cup\{G\in\mathcal{O}; G\subset A\}$. A° の元を A の**内点**という.

2°　(1) $\bar{A}=A\cup A^d$; (2) A が閉$\Longleftrightarrow\bar{A}=A\Longleftrightarrow A^d\subset A$; (3) $x\in\bar{A}\Longleftrightarrow U(x)\cap A\neq\emptyset$ $(\forall U(x)\in\mathfrak{N}(x))$.

証明　$x\notin\bar{A}\Rightarrow U(x)\in\mathfrak{N}(x)$ が存在して $\emptyset=U(x)\cap A\supset(U(x)\setminus\{x\})\cap A\Rightarrow x\notin A^d$, すなわち，$\bar{A}\supset A\cup A^d$. 逆の$\subset$も同様，このことと定義から他の関係式は容易である. ∎

註　$\emptyset^a=\emptyset^d=\emptyset^\circ=\emptyset$, $X^a=X^\circ=X$, $\emptyset^c=X$, $X^c=\emptyset$.

3°　(1)　$A\subset\bar{A}$; $A\subset B\Rightarrow\bar{A}\subset\bar{B}$; $\bar{A}=\bar{\bar{A}}$; $\overline{A\cup B}=\bar{A}\cup\bar{B}$.

(2)　$A^c\cap A=\emptyset$; $A\subset B\Rightarrow A^c\supset B^c$; $A^\circ=A^{cac}$; $A^a=A^{c\circ c}$.

(3)　$A^\circ\subset A$; $A\subset B\Rightarrow A^\circ\subset B^\circ$; $A^\circ=A^{\circ\circ}$; $(A\cap B)^\circ=A^\circ\cap B^\circ$.

ここで，$\bar{\bar{A}}=(A^-)^-$, $A^c=X\setminus A$(補集合)，$A^{ca}=(A^c)^a$, etc. とする.

証明　(1)の最後の等式：$A\cup B\subset\bar{A}\cup\bar{B}$ から $\overline{A\cup B}\subset\bar{A}\cup\bar{B}$, $A\subset A\cup B$ から $\bar{A}\subset\overline{A\cup B}$, 同じく $\bar{B}\subset\overline{A\cup B}$ (i. e., $\bar{A}\cup\bar{B}\subset\overline{A\cup B}$). (2) $A^\circ=A^{cac}$：$A^{ca}\supset A^c\Rightarrow A^{cac}\subset A^{cc}=A$; $G\subset A$, $G\in\mathcal{O}\Rightarrow G^c\supset A^c\Rightarrow G^c=G^{ca}\supset A^{ca}\Rightarrow G=G^{cac}\subset A^{cac}$; ゆえに $A^{cac}=A^\circ$. (3)の最終等式：$(A\cap B)^\circ=(A\cap B)^{cac}=(A^c\cup B^c)^{ac}=(A^{ca}\cup B^{ca})^c=A^{cac}\cap B^{cac}=A^\circ\cap B^\circ$. ∎

以上のほかに，さらに必要な位相的性質の定義を述べよう：x が A の**境界点**$\overset{\text{def}}{\Longleftrightarrow}x\in\bar{A}\cap\overline{A^c}$; x が A の**孤立点**$\overset{\text{def}}{\Longleftrightarrow}x\in A\setminus A^d$; A が X で**稠密** (dense)$\overset{\text{def}}{\Longleftrightarrow}\bar{A}=X$; A が**自己稠密**$\overset{\text{def}}{\Longleftrightarrow}A\subset A^d$; Aが**分散集合**$\overset{\text{def}}{\Longleftrightarrow}A$ は空でない自己稠密な集合を含まない; A が**粗集合**$\overset{\text{def}}{\Longleftrightarrow}\bar{A}^\circ=\emptyset$; 位相空間 X は**第一類**$\overset{\text{def}}{\Longleftrightarrow}X=\bigcup_{n=1}^{\infty}A_n$, A_n は粗集合; X は**第二類**$\overset{\text{def}}{\Longleftrightarrow}X$ が第一類でない; X は**可分**$\overset{\text{def}}{\Longleftrightarrow}$可算集合 $A\subset X$ が存在して $\bar{A}=X$. これらのうちで頻繁に用いられるのは稠密，可分，粗集合，それに第二類集合などである.

註　X の部分集合全体の族を定義域とし，同じく部分集合族を値域とする集合函数 $A\rightarrow A^p$ において p を $a, b, d, \circ$ などの位相的演算を代数的演算の形式に公式化すると，たとえば次のようになる(記号 $1, p\vee q$ などは任意の $A\subset X$ で $A^1=A$, $A^{p\vee q}=A^p\vee A^q$

を表す): $a=1\vee d, aa=a$, $cc=1$, $\circ\circ=\circ$, $a=c\circ c$, $\circ=cac$; さらに, これらを組み合わせると種々な位相演算による等式を得る. これによって, 位相の演算を記号の代数演算として論ずることができる.

2.5 部分空間, 相対位相

位相空間 X の近傍系 $\mathfrak{N}(x)$, $x\in X$ とそれによる位相を $\mathcal{O}$ とするとき, 次が成立:

1° 部分集合 Y に対して, $\mathfrak{N}_Y(x)=\{U(x)\cap Y; U(x)\in\mathfrak{N}(x)\}$, $\mathcal{O}_Y=\{G\cap Y; G\in\mathcal{O}\}$ とおくと, $\mathfrak{N}_Y(y)$, $y\in Y$, は Y における近傍系, $\mathcal{O}_Y$ はそれによる Y の位相となる. この Y を**部分空間**, Y の位相 $\mathcal{O}_Y$ を $\mathcal{O}$ から導かれる**相対位相**という.

2° 上述の X, Y において, それぞれの閉集合全体の族を $\mathcal{F}, \mathcal{F}_Y$ とすれば, $\mathcal{F}_Y=\{F\cap Y; F\in\mathcal{F}\}$. したがって, $\bar{A}^Y=A\cap Y$ は部分空間 Y における A の閉包である.

略証 前半は各 $G\in\mathcal{O}$ に対して $x\in(G\cap Y)^c\cap Y \Longleftrightarrow x\in G^c\cup Y^c$ かつ $x\in Y \Longleftrightarrow x\in G^c\cap Y$, すなわち, $(G\cap Y)^c\cap Y=G^c\cap Y$ から明白. 閉包の方は, $\bar{A}^Y=\cap\{F\cap Y; A\cap Y\subset F\cap Y, F\in\mathcal{F}\}=\cap\{F\cap Y; A\subset F, F\in\mathcal{F}\}$ からいえる. ▌

2.6 収　　束

位相空間 X において, 点列 $\{x_n\}_{n=1}^{\infty}\subset X$ が $x\in X$ に収束するとは, 任意の $U(x)\in\mathfrak{N}(x)$ に対して自然数 n_0 が存在して, $x_n\in U(x)$, $\forall n\geq n_0$, となる場合をいい, $x_n\to x$ で表す. なお, J をある有向集合し, $\alpha\in J$ を添字にもつ X の集合 $\{x_\alpha\}=\{x_\alpha; \alpha\in J\}$ を X のネット (有向系) という. ネット $\{x_\alpha\}$ が $x\in X$ に収束するとは, 任意の $U(x)$ に対して, $\alpha_0\in J$ が存在して, $x_\alpha\in U(x)$, $\forall\alpha\geq\alpha_0(\alpha\in J)$ のときをいう. このとき $x_\alpha\to x$, $x=\lim x_\alpha$ などで表し, x を $\{x_\alpha\}$ の**極限** (点) という.

1° $x\in X$ が $A\subset X$ の集積点 (i. e., $x\in A^d$) $\Longleftrightarrow x$ に収束するネット $\{x_\alpha\}\subset A, x\neq x_\alpha$ ($\forall\alpha\in J$) が存在する.

略証 ($\Rightarrow$): $J\equiv\mathfrak{N}(x)\equiv\{\alpha,\cdots\}$ とおくと, $\forall\alpha\in J$ に対して $A_\alpha\equiv(\alpha\setminus\{x\})\cap A\neq\emptyset$ より元 $x_\alpha\in A_\alpha$ をとれる. $\{x_\alpha\}$ はネットであり, $x_\alpha\to x$ は明白. ($\Leftarrow$) は明らか. ▌

2.7 連続写像と同相写像

2つの位相空間 X, Y 間の写像を φ: $X\to Y$ とし, 任意の $\varphi(x)\in Y$ の近傍 $V(\varphi(x))\in\mathfrak{N}(\varphi(x))$ に対して $U(x)\in\mathfrak{N}(x)$ が存在して $\varphi(U(x))\subset V(\varphi(x))$ となるとき, φ は $x\in X$ で連続であるという. すべての $x\in X$ で φ が連続であるとき, φ を ($X\to Y$ なる) **連続写像**という. 写像 ψ: $X\to Y$ が, X の任意の開集合 G (閉集合 F) に対して $\psi(G)$ ($\psi(F)$) が Y の開 (閉) 集合のとき ψ を**開** (**閉**) **写像**という.

2° 写像 $\varphi(:X\to Y)$ に対して次の4条件は同値である.

(1) φ は連続写像. (2) $x_\alpha\to x\Rightarrow\varphi(x_\alpha)\to\varphi(x)$. (3) $\varphi(\bar{A})\subset\overline{\varphi(A)}$. (4) φ^{-1} は開写像. (5) φ^{-1} は閉写像. ここで (2) は任意のネット $\{x_\alpha\}$ に対して, (3) は任意の

$A \subset X$ に対して成立，を意味する．

（2）⇒（1）**の証明**：（2）かつ（1）が不成立ならば，ある $x \in X$ と $V(\varphi(x)) \in \mathfrak{N}(\varphi(x))$ が存在して $\varphi(\alpha) \not\subset V(\varphi(x))$，$\forall \alpha \in J$ $(J \equiv \mathfrak{N}(x) \equiv \{\alpha, \cdots\})$．ゆえに $y_\alpha = \varphi(x_\alpha) \in \varphi(\alpha) \backslash V(\varphi(x))$ であるネット $\{x_\alpha\} \subset X$ がとれる．これは $x_\alpha \to x$．したがって $\varphi(x_\alpha) \to \varphi(x) \in V(\varphi(x))$．これは矛盾である．（1）⇒（2）：$x_\alpha \to x$ とする．$V(\varphi(x))$ に対して $U(x)$ が存在して $\varphi(U(x)) \subset V(\varphi(x))$．また，ある α_0 で $x_\alpha \in U(x)$，$\forall \alpha \geq \alpha_0$．ゆえに $\varphi(x_\alpha) \in \varphi(U(x)) \subset V(\varphi(x))$，$\forall \alpha \geq \alpha_0$ (i. e., $\varphi(x_\alpha) \to \varphi(x)$)．他の証明もほぼ同様．▌

X から Y の上（全射という）への1対1な写像 φ が連続かつ φ^{-1} も連続であるとき，φ を**同相写像** (homeomorphism)，X と Y は**同位相**または**位相同型**であるといい，これを $X \cong Y$ で表す．

2.8 分離公理

位相空間 X の任意の元 x, y $(x \neq y)$ に対して次の各条件を導入する．（1）$\boldsymbol{T_0}$ **分離公理**：$U(x) \in \mathfrak{N}(x)$ が存在して $y \notin U(x)$ または $U(y) \in \mathfrak{N}(y)$ が存在して $x \notin U(y)$．（2）$\boldsymbol{T_1}$ **分離公理**：$U(x) \in \mathfrak{N}(x)$ が存在して $y \notin U(x)$．（3）$\boldsymbol{T_2}$ **分離公理**：$U(x) \in \mathfrak{N}(x)$，$V(y) \in \mathfrak{N}(y)$ が存在して $U(x) \cap V(y) \neq \emptyset$．（4）$\boldsymbol{T_3}$ **分離公理**：$x \in X$ と $x \notin F$ である閉集合に対して $U(x) \in \mathfrak{N}(x)$ と F を含む開集合 $V(F)$ が存在して $U(x) \cap V(F) = \emptyset$．（5）$\boldsymbol{T_3'}$ **分離公理**：（4）と同じ x, F に対し連続函数 f: $X \to [0, 1]$ が存在して $f(x)=0$，$f(F)=1$．（6）$\boldsymbol{T_4}$ **分離公理**：閉集合 $F_1, F_2 (F_1 \cap F_2 = \emptyset)$ に対して，開集合 U_1, U_2 が存在して $F_1 \subset U_1$，$F_2 \subset U_2$ かつ $U_1 \cap U_2 = \emptyset$．

1° 各分離公理間の関係：（1）$T_2 \Rightarrow T_1 \Rightarrow T_0$；$T_0 + T_3 \Longleftrightarrow T_1 + T_3 \Longleftrightarrow T_2 + T_3 \Leftarrow T_0 + T_3' \Leftarrow T_0 + T_4 \Longleftrightarrow T_1 + T_4 \Longleftrightarrow T_2 + T_4$．（2）$T_3 \Longleftrightarrow x \in X$ と $U(x) \in \mathfrak{N}(x)$ に対して $V(x) \in \mathfrak{N}(x)$ が存在して $\overline{V(x)} \subset U(x)$．（3）**(Urysohn の補題)** $T_4 \Longleftrightarrow$ 互いに素な閉集合 F_1, F_2 に対して連続函数 f: $X \to [0, 1]$ が存在して $f(F_1)=0$，$f(F_2)=1$．

これらの証明は（2）（これは容易）と（3）をまず示し，しかる後に（1）を証明する．

$T_0, T_1, T_2, T_0+T_3, T_0+T_3', T_0+T_4$ の各分離公理を満たす位相空間を各々 $\boldsymbol{T_0}$，$\boldsymbol{T_1}$，$\boldsymbol{T_2}$ または **Hausdorff**，**正則**，**完全正則**，**正規空間**と名付ける．これらのうちでとくに重要なのは，Hausdorff 空間と完全正則空間，それに正規空間である．

2.9 コンパクト性

位相空間 X の**開被覆** (Open covering) $\mathcal{U} \equiv \{U_\alpha; \alpha \in J\}$ (i. e., $U_\alpha \in \mathcal{O}$, $\bigcup U_\alpha = X$) から**有限部分被覆** $\{U_{\alpha_k}; k=1, 2, \cdots, n\}$ が選べるとき X は**コンパクト**であるという．これは X の**有限交差性**をもつ任意の閉集合族 $\mathcal{F}_0 = \{F_\alpha; \alpha \in J\}$（すなわち，任意の有限部分族 $\{F_{\alpha_k}; k=1, 2, \cdots, n\} \subset \mathcal{F}_0$ に対して $\bigcap_{k=1}^{n} F_{\alpha_k} \neq \emptyset$）はつねに $\bigcap F_\alpha \neq \emptyset$ であることと同値である．

$A \subset X$ が相対位相に関してコンパクトのとき A を単に**コンパクト**といい，また $\bar{A}$ が

コンパクトのとき**相対コンパクト**という．X のすべての元が閉包をとればコンパクトになる（このとき“**コンパクト閉包な**”という）近傍をもつとき，X を**局所コンパクト**であるという．

$A(\subset X)$ の無限部分集合 B が少なくともひとつ集積点をもつとき，A を**可算コンパクト**であるという．また，A の点列（＝可算）がつねに収束部分列を含むとき，A を**点列コンパクト**という．

1° 位相空間 X におけるコンパクト性の定義で，開被覆 $\{U_\alpha;\alpha\in J\}$ と有限交差性をもつ閉集合族 $\{F_\alpha;\alpha\in J\}$ の J を自然数の全体 N に置き換えたときの条件は，ともにそれぞれ可算コンパクト性と同値となる．X が第二可算公理を満たせばコンパクト$\Longleftrightarrow$可算コンパクトである．

2° （1）X がコンパクト⇨閉部分集合はコンパクト；（2）X が Hausdorff 空間のときコンパクト部分集合は閉．

3° コンパクト X から Hausdorff 空間 Y の上への1対1連続写像 φ は同相写像であり，Y もコンパクトである．

2° の証明は容易であり，3° の証明はコンパクト X の連続写像による像はコンパクトであることと 2° の後半を用いる．

4° コンパクト Hausdorff 空間 X は正規であり，X 上の実数値連続函数 f は最大値，最小値をもつ．

略証 これの後半は容易．前半は，閉集合 F_1, F_2 が $F_1\cap F_2=\emptyset$ ならば，ともにコンパクトから，任意の $x\in F_1$, $y\in F_2$ は $U(x)\cap V(y)=\emptyset$ である近傍 $U(x), V(y)$ をもつ，コンパクト性を被覆 $\{U(x);x\in F_1\}$, $\{V(y);y\in F_2\}$ に適応して開集合 $F_1\subset U(F_1)$, $F_2\subset V(F_2)$ がつくれて $U(F_1)\cap U(F_2)=\emptyset$. ▌

5° 局所コンパクト Hausdorff 空間は完全正則である．

2.10 コンパクト化と Tychonoff の定理

位相空間 X に対して，コンパクト空間 $\tilde{X}$ と X から $\tilde{X}$ の稠密な部分空間への同相写像 φ が存在するとき，対 $(\tilde{X},\varphi)$ を X の**コンパクト化**という．

1° **一点コンパクト化**：局所コンパクト Hausdorff 空間 X に対するコンパクト化の空間 $\tilde{X}$ は X に一点（∞とかく）を付加して構成される．このとき∞の近傍系は $\mathfrak{N}(\infty)=\{\tilde{X}\backslash F;F\subset X$ はコンパクト$\}$ である．この証明は容易．

2° **Čech のコンパクト化**：完全正則空間 X に対するコンパクト化空間 $\tilde{X}=\beta(X)$ は Hausdorff 空間（したがって正規空間）で X との有界連続函数 f は一意に $\beta(X)$ 上に拡張される．この空間 $\beta(X)$ は位相同型の範囲で一意に定まる（X と $\tilde{X}$ 上の有界連続函数空間 $C(X)$ と $C(\tilde{X})$ とは等距離同型となる）．この証明は，Banach 代数（第2巻参照）の表現を用いる函数解析的方法で行うことにより比較的容易に示される．

3° **Tychonoff の定理**：$\{(X_\alpha,\mathcal{O}_\alpha);\alpha\in J\}$ を位相空間の族とするとき $\Pi X_\alpha\equiv\prod_{\alpha\in J}X_\alpha$

の（Tychonoff の）直積位相は筒集合の族 $\{U_{\alpha_1}\times\cdots\times U_{\alpha_n}\times\prod_{\beta\in J}X_\beta;\ \beta\neq\alpha_i,\ U_{\alpha_j}\in\mathcal{O}_{\alpha_j},\ j=1,2,\cdots,n<\infty\}$ を基とする位相である．（1）この位相は射影 $P_\alpha\colon\prod_{\alpha\in J}X_\alpha\to X_\alpha$ を連続にする最も弱い位相である．（2）各 $X_\alpha(\alpha\in J)$ が Hausdorff 空間ならば $\prod X_\alpha$ もそうである．（3）各 X_α $(\alpha\in J)$ がコンパクト$\Longleftrightarrow\prod X_\alpha$ がコンパクト．

2.11　連結性，不連結性

位相空間 X が**連結**であるとは，$G_1\cap G_2=\emptyset$ であるいかなる開集合 G_1, G_2 に対しても $X\neq G_1\cup G_2$ のときをいう．もし $X=G_1\cup G_2$ と表せるとき，G_1, G_2 はたがいに補集合となるのでともに閉集合である．したがって，X が連結$\Longleftrightarrow X$ の閉かつ開な集合は X と $\emptyset$ のみ．X の閉かつ開な集合の全体が位相基をなすとき，X を**完全（totally）不連結**という．X がコンパクトHausdorff 空間のとき，任意の開集合の閉包が開集合であるときXを**最端（extremal）不連結**であるという（**Stone 空間**ともいう），不連結$\Leftarrow$完全不連結$\Leftarrow$最端不連結である（なお，X が連結でない場合を不連結という）．

2.12　距離空間

［**定義 2.1**］ 集合 X の任意の対 $x, y\in X$ に実数 $d(x,y)$ が対応し，次の条件を満足するとき，(X,d) または単に X を**距離空間**という．

(**d.1**)　$d(x,y)\geq 0$; $d(x,y)=0\Longleftrightarrow x=y$　（分離性）．

(**d.2**)　$d(x,y)=d(y,x)$　（対称性）．

(**d.3**)　$d(x,y)\leq d(x,z)+d(z,y)$　（三角不等式）．

この d を X 上の**距離（函数）**という．$B_r(x)\equiv\{y\in X; d(x,y)<r\}$ を**中心 x，半径 r の開球**という．これは近傍系をなす：$\mathfrak{N}(x)=\{B_r(x); r>0\}$．これにより X は位相空間となり，$\mathcal{B}\equiv\{B_r(x); x\in X, r>0\}$ は X の位相基となる．点列 $\{x_n\}\subset X$ が（可算）ネットとして元 $x\in X$ に収束することと $d(x_n,x)\to 0$ は同値である（$x_n\to x$ で表す）．点列 $\{x_n\}$ が $d(x_n,x_m)\to 0$ $(m,n\to 0)$ であるとき，$\{x_n\}$ を **Cauchy 列という**．Cauchy 列がつねに収束するとき X は**完備**であるという．任意の $\varepsilon>0$ に対して，有限集合 $\{x_1,\cdots,x_n\}\subset X$ が存在して $X=\cup B_\varepsilon(x_i)$ であるとき X を**全有界**という．部分集合 $A\subset X$ に対しても，距離函数 d を A の上に制限することによって部分距離空間として上述の定義・概念はすべて適用される．距離空間 X において次の命題が成り立つ：

1°（1） X は第一可算公理を満たす正規空間である．

（2） X が全有界$\Rightarrow$可分$\Longleftrightarrow$第二可算性．

（3） コンパクト＝点列コンパクト＝可算コンパクト＝全有界＋完備．

（4） $A\subset X$ が部分距離空間として完備$\Longleftrightarrow A$ が X で閉．

（5） **Baire カテゴリ定理**：完備距離空間は第二類である．

略証：X が全有界$\Rightarrow$可分：各 n に対し，有限集合 $x_1^{(n)},\cdots,x_{k(n)}^{(n)}\in X$ が存在して $X=\bigcup_{j=1}^{k(n)}B_{1/n}(x_i)$．集合 $\{x_j^{(n)}|j=1,2,\cdots,k(n);\ n=1,2,\cdots\}$ は可算かつ X で稠密であ

る．これと第二可算性との同値性は容易．これを用いて可算コンパクト⇨コンパクトがすぐいえる．（4）：X は粗集合の列 $\{A_n\}$ によって $X=\cup A_n$ と表されないことを示す．もし，$X=\cup A_n$ であるとする．まず半径1の閉球 $B_0=\{x;d(x_0,x)\leq 2\}$ をひとつ定める．A_1 が粗集合であるから半径 ≤ 1 の閉球 $B_1\subset B_0$ をとって $A_1\cap B_1=\emptyset$ とできる．A_2 は同じ粗集合であるから半径 $\leq 1/2$ の閉球 $B_2\subset B_1$ をとし $A_2\cap B_2=\emptyset$ とできる．以下これをくり返し，単調減少な閉球列 $\{B_n\}$ を得る．X が完備であるから $\cap B_n\neq\emptyset$．ゆえに $x\in\cap B_n$ が存在する．一方 $\cap B_n\subset(\cup A_m)^c=X^c=\emptyset$，これは矛盾．▌

註 1　距離空間 X の部分集合 $B\subset X$ の直径を $D(B)=\sup\{d(x,y);x,y\in B\}$ で定義する．X が完備 $\Longleftrightarrow X$ の閉集合列 $S_n\subset X$，$S_1\supset S_2\cdots$，$D(S_n)\downarrow 0$ はつねに $\cap S_n\neq\emptyset$．これの証明は容易である．（4）の証明はこれを用いた．

註 2　1°の（2）や上述の註1は第4章4.3節の ε-エントロピーの議論で用いる．

2.13 距離空間の完備化，その他の諸定理

(X,d) を完備でない距離空間とする．$\mathcal{C}$ を X の Cauchy 列の全体とし，各対 $\{x_n\}$，$\{y_n\}\in\mathcal{C}$ に対して $\{x_n\}\sim\{y_n\}\overset{\text{def}}{\Longleftrightarrow}d(x_n,y_n)\to 0$ $(n\to\infty)$ と～を定めると，～は $\mathcal{C}$ の同値関係となる．これを用いて商空間を定義する：$\tilde{X}=\mathcal{C}/\sim=\{\tilde{x},\tilde{y},\cdots\}$ $(\tilde{x}=[\{x_n\}]$，$\tilde{y}=[\{y_n\}],\cdots)$ とおき，$\tilde{x},\tilde{y}\in\tilde{X}$ に対して $\tilde{d}(\tilde{x},\tilde{y})=\lim d(x_n,y_n)$ と $\tilde{d}$ を定義すると，

1°　$(\tilde{X},\tilde{d})$ は完備な距離空間である．

略証　$(\tilde{X},\tilde{d})$ が距離空間となることは明白．完備性は $\{\tilde{x}(k)\}$ を $(\tilde{X},\tilde{d})$ の Cauchy 列とする．$\tilde{x}(k)\equiv(\{x(k,n)\}_{n=1}^{\infty})$ として，$x_k\in\{x(k,n)\}$ $(k=1,2,\cdots)$ を取り出し，$|d(x_j,x_k)-\tilde{d}(\tilde{x}(j),\tilde{x}(k))|<\varepsilon/(j+k)$ とできる．すると $(\tilde{x}\equiv)\{x_k\}\in\mathcal{C}$ がいえ，$\tilde{d}(\tilde{x}(k),\tilde{x})\to 0$ $(k\to\infty)$ となる．各 $x\in X$ に対し，$x_n\equiv x$ $(\forall n)$ とおくと，一点集合は $\{x\}\in\mathcal{C}$ と見なせる．対応 φ: $x\in X\to\tilde{x}=[\{x_n\}]$ $(x_n\equiv x)$ は $X\to\varphi(X)\subset\tilde{X}$ となる等距離写像 $(d(x,y)=d(\varphi(x),\varphi(y)))$ で $\varphi(X)$ は $\tilde{X}$ で稠密 (i.e., $\overline{\varphi(X)}=X$) である．これらの証明は $\tilde{X}$ の構成から明らかである．▌

一般に位相空間 X 上に距離函数 d を適当に定めて，d による位相が X にあらかじめ与えられていた位相と同位相になるとき，X は**距離化可能**であるという．

2°　**Urysohn の定理**：第二可算公理を満たす正則空間は距離化可能である．したがって，局所コンパクトな完全正則空間またはコンパクト Hausdorff 空間が第二可算公理を満たすとき，それらはともに距離化可能である（2.9 の 4° と 5° 参照）．

最後に**距離空間上の連続写像について**の有用な定理を述べよう．(X,d_1)，(Y,d_2) を2つの距離空間とする．f: $X\to Y$ である写像（Y-値函数）f の族を $\boldsymbol{F}$ とする．

3°　f が $x\in X$ で連続 $\Longleftrightarrow$ 任意の $\varepsilon>0$ に対して $\delta=\delta(f,x,\varepsilon)>0$ が存在して $d_1(x,x')<\delta$ ならば $d_2(f(x),f(x'))<\varepsilon$．

距離空間上では実数直線 $\boldsymbol{R}$ 上と同様強い連続性の概念が定義される：f が X 上で**一様連続** $\overset{\text{def}}{\Longleftrightarrow}$ 任意の $\varepsilon>0$ に対して $\delta=\delta(f,\varepsilon)>0$ が存在して，$d_1(x,x')<\delta$ ならばつねに

$d_2(f(x),\ f(x'))<\varepsilon$; 函数族 $\boldsymbol{F}$ が**同程度連続**$\overset{\text{def}}{\Longleftrightarrow}$任意の $\varepsilon>0$ と $\forall x\in X$ に対して, $\delta=\delta(\varepsilon, x)>0$ が存在して $d_1(x, x')<\delta$ ならば $d_2(f(x), f(x'))<\varepsilon$ ($\forall f\in\boldsymbol{F}$), ここで $\delta=\delta(\varepsilon)$ ($x\in X$ に依存せず) のとき, これは**同程度一様連続**の定義となる. これらの連続性を用いて次の定理が示される.

4° 函数列 $\boldsymbol{F}=\{f_n\}_{n=1}^{\infty}$ が同程度連続であるとする, このとき, (1) $f_n(x)\to f(x)$ ($\forall x\in X$)$\Longrightarrow f$ は連続; (2) Y が完備で, $f_n(x)\to f(x)$ ($\forall x\in D,\ \bar{D}=X$)$\Longrightarrow f_n(x)\to f(x)$ ($\forall x\in X$).

5° **Ascoli-Arzelá の定理**: $X=[a, b], Y=\boldsymbol{R}$ のとき, $\boldsymbol{F}$ が同程度連続で一様有界ならば $\boldsymbol{F}$ は一様収束する部分列を含む.

3. 測度と積分

区間 $[a, b]$ 上の函数 f が Riemann 可積分であるとは, $[a, b]$ の任意の分割 Δ: $a=x_0<x_1<\cdots<x_n=b$ に対して, $S(\Delta)=\sum_{j=1}^{n} f(t_j)\ (x_j-x_{j-1})\ (x_{j-1}\leqq t_j<x_j)$ とおくと, $|\Delta|\equiv\max_j(x_j-x_{j-1})\to 0$ のとき $S(\Delta)$ が t_j のとり方に無関係にある有限確定値に収束する場合をいうが, 今 f として, $f(x)=0$ ($x\in\boldsymbol{Q}$), $=1$ ($x\in\boldsymbol{R}\backslash\boldsymbol{Q}$) をとると, 任意の分割 Δ に対して, $t_j\in\boldsymbol{Q}$ とすれば $S(\Delta)=0$, $t_j\in\boldsymbol{R}\backslash\boldsymbol{Q}$ とすれば $S(\Delta)=b-a$ となり, この f は Riemann 可積分ではなくなってしまう. しかし, 区間 $[a, b]$ における有理数の個数は無理数の個数に比べてはるかに少ないので, 上記の f の積分値は $\int_a^b f(x)dx=b-a$ とする方がもっともらしい. この種の疑問に直面し Lebesgue はものを“測る”という概念それ自体の批判的検討を行い, 新たな測度論・積分論を確立したのである.

この章では抽象的な集合上の測度・積分論の重要な結果について概説し, Lebesgue 測度論, 測度論的確率論に言及する.

3.1 測度と測度空間

集合 Ω の空でない部分集合族 $\mathcal{R}$ が

(1) $A, B\in\mathcal{R}\Rightarrow A\backslash B\in\mathcal{R}$, (2) $A, B\in\mathcal{R}\Rightarrow A\cup B\in\mathcal{R}$

を満たすとき, $\mathcal{R}$ は**集合環**であるといい, (2) の代わりに,

(2′) $A_k\in\mathcal{R}(k=1, 2, \cdots)\Rightarrow\bigcup_{k=1}^{\infty} A_k\in\mathcal{R}$

を満たすとき, $\mathcal{R}$ は **σ-集合環** (σ-ring) であるという. さらに, Ω 自身を含む (σ-) 集合環を (σ-) **集合体** (field) または (σ-) **代数** (algebra) とよび, $\mathcal{F}$, または Ω を明記する必要があるときは $\mathcal{F}_\Omega$ で表す. σ-集合体の族 $\{\mathcal{F}_\alpha;\alpha\in J\}$ において, $\bigcap_{\alpha\in J}\mathcal{F}_\alpha$ も σ-集合体になり, このことから, Ω の任意の部分集合族 $\mathcal{E}$ が σ-集合体になっていなくても, この $\mathcal{E}$ を含む最小の σ-集合体を構成することができる. この σ-集合体を $\mathcal{E}$ **によって生成された σ-集合体**といい, $\boldsymbol{\sigma}(\mathcal{E})$ で表す.

集合 X の部分集合族 $\mathcal{E}$ において, $\{A_n\}\subset\mathcal{E}$ が単調列 (i. e., $A_n\subset A_{n+1}$ または $A_n\supset$

A_{n+1}) であれば，$\lim A_n \in \mathcal{E}$ となるとき，$\mathcal{E}$ を Ω の**単調クラス** (monotone class) という．単調クラスの族 $\{\mathcal{E}_\alpha; \alpha \in J\}$ に対して，$\mathcal{E} \equiv \bigcap_{\alpha \in J} \mathcal{E}_\alpha$ も単調クラスとなり，このことより，Ω の任意の部分集合族 $\mathcal{E}$ に対して，$\mathcal{E}$ を含む最小の単調クラスが存在することがわかる．この単調クラスを $m(\mathcal{E})$ で表す．このとき次の重要な定理が証明される．

〈定理 3.1〉 単調クラス定理：$\mathcal{E}$ が集合体のとき $m(\mathcal{E}) = \sigma(\mathcal{E})$.

集合 Ω のある集合体 $\mathcal{F}$ から $\overline{\boldsymbol{R}} \equiv \boldsymbol{R} \cup \{\pm\infty\}$ への写像 μ を**集合函数** (set function) という（ただし，μ は定義されること，すなわち，$\infty - \infty$ とならないことが必要!）．この μ が，$A \cap B = \emptyset$ となる任意の $A, B \in \mathcal{F}$ に対して $\mu(A \cup B) = \mu(A) + \mu(A)$ を満たすとき，μ は（**有限**）**加法的**であるという．この加法的集合函数 μ は次の性質を有する．

1° μ が恒等的に $\pm\infty$ でないとき，$\mu(\emptyset) = 0$.

2° $A \subset B$, $-\infty < \mu(A) < \infty \Rightarrow \mu(B \backslash A) = \mu(B) - \mu(A)$.

3° μ: $\Omega \to \overline{\boldsymbol{R}^+} \equiv \boldsymbol{R}^+ \cup \{+\infty\}$ のとき，$A \subset B \Rightarrow \mu(A) \leqq \mu(B)$.

4° μ が 3° と同じとき，$A_k \in \mathcal{F}(k = 1, 2, \cdots, n) \Rightarrow \mu(\bigcup_{k=1}^{n} A_k) \leqq \sum_{k=1}^{n} \mu(A_k)$.

なお，$A_i \cap A_j = \emptyset$ $(i \neq j)$ で $\{A_n; n \in \boldsymbol{N}\} \subset \mathcal{F}$ かつ $\bigcup_n A_n \in \mathcal{F}$ となる $\{A_n\}$ に対して，$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ を満たすとき，μ は **σ-加法的**あるいは**完全加法的**であるという．

Ω の σ-集合体 $\mathcal{F}$ から $\overline{\boldsymbol{R}}^+$ への σ-加法的集合函数 μ で $\mu(\emptyset) = 0$ を満たすものを Ω 上の**測度** (measure) という．これに対して一般の σ-加法的集合函数を**符号付測度** (signed measure) ということもある．さらに，$\mu(\Omega) < +\infty$ となる測度を**有限測度**，$\mu(\Omega) = 1$ となるものを**確率測度**といい，$\mu(\Omega) = +\infty$ であっても $\Omega = \bigcup_{n=1}^{\infty} A_n$ を満たす $\{A_n\} \subset \mathcal{F}$ が存在して $\mu(A_n) < +\infty$ となるとき，測度 μ は **σ-有限**であるという．

一般に，$\mathcal{F}$ の元を**可測集合**，組 $(\Omega, \mathcal{F})$ を**可測空間**，μ が測度のとき $(\Omega, \mathcal{F}, \mu)$ を**測度空間**とよぶ．

問 1 $\mathcal{F} = 2^{\Omega}$（べき集合）のとき，次の μ は測度になることを示せ．(1) $\mu_\omega(A) = 1$ $(\omega \in A)$, $= 0$ $(\omega \notin A)$. (2) $\mu(A) = (A$ の元の個数). ((1) の測度を **Dirac** 測度，(2) の測度を**数測度** (counting measure) という.)

測度空間 $(\Omega, \mathcal{F}, \mu)$ において，$\mu(A) = 0$, $A \in \mathcal{F}$, $B \subset A$ であれば $B \in \mathcal{F}$ となるとき，μ は**完備** (complete) であるといい，$(\Omega, \mathcal{F}, \mu)$ を**完備な測度空間**という．また $\mu(N) = 0$ となる $N \in \mathcal{F}$ を **μ-零集合** (μ-null set) とよぶ．$(\Omega, \mathcal{F}, \mu)$ が完備でないときは，次のようにして完備な測度空間 $(\Omega, \widetilde{\mathcal{F}}, \tilde{\mu})$ を構成することができる：(1) $\widetilde{\mathcal{F}} \equiv \{A \cup M; A \in \mathcal{F}, M \subset N \in \mathcal{F}, \mu(N) = 0\}$ が σ-集合体となる．(2) $Q \equiv A \cup M \in \widetilde{\mathcal{F}}$ に対して，$\tilde{\mu}(Q) \equiv \mu(A)$ とおくと，$\tilde{\mu}$ は $\widetilde{\mathcal{F}}$ 上の完備な測度で，$\tilde{\mu}(A) = \mu(A)$ $(\forall A \in \mathcal{F})$. このようにして得られた $(\Omega, \widetilde{\mathcal{F}}, \tilde{\mu})$ を $(\Omega, \mathcal{F}, \mu)$ の**完備化**という．

さて，写像 μ^*: $2^\Omega \to \overline{\boldsymbol{R}}^+$ が条件：(i) $\mu^*(\emptyset)=0$; (ii) $A\subset B \Rightarrow \mu^*(A)\leqq\mu^*(B)$; (iii) $\mu^*(\bigcup_{n=1}^{\infty}A_n)\leqq\sum_{n=1}^{\infty}\mu^*(A_n)$; を満たすとき，$\mu^*$ は Ω 上の**外測度**であるという．この μ^* はつねに $\mu^*(B)\leqq\mu^*(A\cap B)+\mu^*(A^c\cap B)$ を満たしているが，とくにすべての $B\in 2^\Omega$ に対して等号が成立するような A を **μ^*-可測集合**という．μ^*-可測な集合の全体 $\mathcal{F}^*$ は σ-集合体となり，$\mathcal{F}^*$ 上で μ^* は測度となる．

問 2　$\mathcal{E}$ を Ω のある集合体とし，μ_0 を $\mathcal{E}$ 上の有限加法的な正値集合函数で $\mu_0(\emptyset)=0$ を満たすものとする．このとき，$\mu^*(Q)\equiv\inf\{\sum_{n=1}^{\infty}\mu_0(A_n);\ A_n\in\mathcal{E},\ Q\subset\bigcup_{n=1}^{\infty}A_n\}$ は Ω 上の外測度となることを示せ．また μ^*-可測集合の全体 $\mathcal{F}^*$ は $\mathcal{E}\subset\mathcal{F}^*$ を満たすことを証明せよ．

今，$\mathcal{E},\mu_0$ を問2のそれと同じものとし，$\mathcal{F}=\sigma(\mathcal{E})$ とする．このとき，

5°　**拡張定理**：μ_0 が $\mathcal{E}$ 上で σ-加法的（すなわち，$\{A_n\}\subset\mathcal{E}$ で $A_n\cap A_m=\emptyset\,(n\neq m)$, $\bigcup_{n=1}^{\infty}A_n\in\mathcal{E}\Rightarrow\mu_0(\bigcup_{n=1}^{\infty}A_n)=\sum_{n=1}^{\infty}\mu_0(A_n)$）であることと，$\mathcal{F}(=\sigma(\mathcal{E}))$ 上の測度 μ が存在して，$\mu\restriction\mathcal{E}=\mu_0$ となることとは同値（μ_0 が $\mathcal{E}$ 上で σ-有限のときこの拡張は一意）．

ところで，$\mathcal{E}\equiv\{\bigcup_{k=1}^{n}[a_k,b_k);-\infty<a_1<b_1<\cdots<a_n<b_n\leqq+\infty\}$ および $\mu_0(\bigcup_{k=1}^{n}[a_k,b_k))=\sum_{k=1}^{n}(F(b_k)-F(a_k))$（$F$ は $\boldsymbol{R}$ 上の左連続な単調増加函数）に対して，5° を適用したとき得られる測度空間 $(\Omega,\mathcal{F},\mu)$ が **Lebesgue-Stieltjes 測度空間**，とくに $F(t)=t$ の場合が，**Lebesgue 測度空間**である．なお，$F(t)=t$ のとき，問2の仕方で $\mathcal{E},\mu_0$ より構成した $\mathcal{F}^*$ の元を **Lebesgue 可測集合**とよぶ．

3.2　位相空間上の測度

ここでは集合 Ω を局所コンパクトな Hausdorff 空間とし，その上での測度について考察する．

Ω のすべての開集合から生成される σ-集合体を **Borel 集合体**といい，$\mathcal{B}$ または $\mathcal{B}_\Omega$ で表す．また，$\mathcal{B}$ の元を **Borel 集合**という．

$\{G_n\},\{F_n\}$ を Ω の開集合および閉集合の可算列とするとき，$\bigcap_{n=1}^{\infty}G_n$, $\bigcup_{n=1}^{\infty}F_n$ を各々 **G_δ-集合**，**F_σ-集合**という．コンパクトな G_δ-集合から生成される σ-集合体を **Baire 集合体**といい，$\mathcal{B}_0$ で表す．$\mathcal{B}_0$ の元を **Baire 集合**とよぶ．

1°　$\mathcal{B}_0\subset\mathcal{B}$.

2°　すべてのコンパクトな Baire 集合は G_δ-集合．

3°　Ω が距離空間ならばすべての開集合は G_δ-集合であり，したがって，$\mathcal{B}_0=\mathcal{B}$.

Ω 上の **Borel (Baire) 測度**とは，任意のコンパクト集合（コンパクト G_δ 集合）$A\subset\Omega$ において，$\mu(A)<\infty$ となる $\mathcal{B}(\mathcal{B}_0)$ 上の測度のことである．

Borel 測度 μ が**正則**であるとは，任意の $A\in\mathcal{B}$ に対して

$$\mu(A)=\inf\{\mu(O);\ A\subset O,\ O\in\mathcal{B}\text{ で開集合}\}$$

$$=\sup\{\mu(C);\ C\subset A,\ C\in\mathscr{B}\ \text{でコンパクト}\}$$

が成立する場合をいう．なお，すべての Baire 測度は上の意味で正則であり，Baire 測度と正則な Borel 測度との間には 1 対 1 の対応が存在する．

3.3 可測函数

可測空間 $(\Omega, \mathscr{F})$ から他の可測空間 $(\Omega', \mathscr{F}')$ への写像 f がすべての $A\in\mathscr{F}'$ に対して $f^{-1}(A)\in\mathscr{F}$ であるとき，f は可測であるという．とくに，$\Omega'=\overline{\boldsymbol{R}}$ か $\overline{\boldsymbol{C}}$ のとき，f は $\mathscr{F}$ に関する**可測函数**であるという．以下，$\mathscr{F}$ に関する可測函数の全体を $\mathfrak{M}_{\mathscr{F}}$ で表し，$f, g, f_n (n\in\boldsymbol{N})\in\mathfrak{M}_{\mathscr{F}}$ とする．

1° $\{\omega\in\Omega;\ a<f(\omega)<b,\ a, b\in\boldsymbol{R}\}\in\mathscr{F}$.

2° $f+g$, fg, $|f|$, αf $(\alpha\in\boldsymbol{C})$, f/g, $\sup f_n$, $\inf f_n$, $\overline{\lim} f_n$, $\underline{\lim} f_n$, $\lim f_n$ は可測函数．

ところで，測度空間 $(\Omega, \mathscr{F}, \mu)$ において，$\omega\in\Omega$ に関するある命題 $P(\omega)$ が**ほとんど到るところ** (almost everywhere, **a. e.** と略) で成立するとは μ-零集合 N を除いて $P(\omega)$ が成立する場合をいい，"$\boldsymbol{P(\omega)}$ **$\boldsymbol{\mu}$-a. e.**" とかく．たとえば，2 つの可測函数 f, g がほとんど到るところで等しいとは，$\mu\{\omega\in\Omega;\ f(\omega)\neq g(\omega)\}=0$ のことであり，この場合 $\boldsymbol{f=g(\mu}$**-a. e.)** とかく．

とくに次の可測函数は重要である：任意の $A\in\mathscr{F}$ に対して $1_A(\omega)=1$ $(\forall\omega\in A)$, $=0$ $(\forall\omega\notin A)$ となる函数 $1_A(\omega)$ を A に関する**特性函数**とよび，たがいに素な有限個の $\{A_k\}_{k=1}^n\subset\mathscr{F}$ を用いて $\sum_{k=1}^n\alpha_k 1_{A_k}(\omega)$ $(\alpha_k\in\boldsymbol{R}$ または $\alpha_k\in\boldsymbol{C})$ と表される函数を**単函数** (simple function) という．

さて，ここで可測函数列の収束を論じておこう．以下，$(X, \mathscr{F}, \mu)$ を測度空間，$\{f_n\}$ を可測函数の列とする．

[定義 3.1]

(1) **概収束**：$f_n\xrightarrow{\text{a.e.}}f \overset{\text{def}}{\Longleftrightarrow} \mu$-零集合の元を除いて，$f_n(\omega)\to f(\omega)$.

(2) **測度収束**：$f_n\xrightarrow{\mu}f \overset{\text{def}}{\Longleftrightarrow}$ 任意の $\varepsilon>0$ に対して，$\mu(\{\omega\in\Omega;\ |f_n(\omega)-f(\omega)|\geqq\varepsilon\})\to 0$ $(n\to\infty)$.

(3) **一様収束**：$f_n\xrightarrow{u}f \overset{\text{def}}{\Longleftrightarrow} \|f_n-f\|\equiv\sup\{|f_n(\omega)-f(\omega)|;\ \omega\in\Omega\}\to 0$.

(4) **概一様収束**：$f_n\xrightarrow{\text{a.u.}}f \overset{\text{def}}{\Longleftrightarrow}$ 任意の $\varepsilon>0$ に対して，$A\in\mathscr{F}$ が存在して $\mu(A)<\varepsilon$ かつ $A^c\equiv\Omega\backslash A$ 上で $f_n\xrightarrow{u}f$.

収束についてのこれらの概念の相関を示しておこう．以下，f_n, f はほとんど到るところで有限値 (i. e., $|f_n(\omega)|<\infty$, $|f(\omega)|<\infty$, μ-a. e.) をとる可測函数とする．

3° $\mu(\Omega)<\infty$ のとき，$f_n\xrightarrow{\text{a.e.}}f$ ならば $\mu(\bigcup_{k=n}^{\infty}\{\omega\in\Omega;\ |f_k(\omega)-f(\omega)|\geqq\varepsilon\})\to 0$ $(n\to\infty)$.

4° **Egoroff の定理**：$\mu(\Omega)<\infty$ のとき，$f_n\xrightarrow{\text{a.e.}}f$ と $f_n\xrightarrow{\text{a.u.}}f$ は同値．

5° $f_n \xrightarrow{\mu} f$ のとき，$\{f_n\}$ の任意の部分列 $\{f_{n_k}\}$ は $f_{n_{k_i}} \xrightarrow{\text{a.e.}} f$ $(i\to\infty)$ となる列 $\{f_{n_{k_i}}\}$ を含む．

6° $\mu(\Omega)<\infty$ のとき，$f_n \xrightarrow{\text{a.e.}} f$ ならば $f_n \xrightarrow{\mu} f$ である．

7° $f_n \xrightarrow{\text{a.u.}} f$ ならば $f_n \xrightarrow{\text{a.e.}} f$ かつ $f_n \xrightarrow{\mu} f$.

8° $f_n \xrightarrow{\text{a.e.}} f$ ならば $f_{n_k} \xrightarrow{\text{a.u.}} f$ となる $\{f_n\}$ の部分列 $\{f_{n_k}\}$ が存在する．

〈定理 3.2〉 任意の可測函数 f に対して，単函数の列 $\{\xi_n\}$ が存在して，$\xi_n \xrightarrow{\text{a.e.}} f$.

略証 $f=f_+-f_-$ $(f_+\equiv\max\{f,0\},\ f_-\equiv(-f)_+)$ とかけるから，$f\geqq 0$ の場合に証明すれば十分．今，$E_n^{(k)}\equiv\left\{\omega\in\Omega;\ \frac{k-1}{2^n}\leqq f(\omega)<\frac{k}{2^n}\right\}$ $(k=1,2,\cdots,n\cdot 2^n)$, $E_n^{(n\cdot 2^n+1)}\equiv\{\omega\in\Omega;\ f(\omega)\geqq n\}$ とおくと，$\{E_n^{(k)}\}$ はたがいに素で $\bigcup_k E_n^{(k)}=\Omega$ を満たす．ここで $\xi_n(\omega)=\sum_{k=1}^{n\cdot 2^n+1}\frac{k-1}{2^n}\cdot 1_{E_n^{(k)}}(\omega)$ と定めると，$\xi_n\uparrow f$ がわかる．∎

3.4 積分と収束定理

測度空間 $(\Omega,\mathcal{F},\mu)$ 上の単函数 $\xi(\omega)=\sum_{k=1}^{n}\alpha_k 1_{A_k}(\omega)$ の積分を $\int_\Omega \xi(\omega)d\mu(\omega)\left(=\int_\Omega \xi d\mu\ \text{とかく}\right)=\sum_{k=1}^{n}\alpha_k\mu(A_k)$ で定義する．また，任意の正値可測函数 f は〈定理 3.2〉より単函数の増加列 $\{\xi_n\}$ で近似できる (i.e., $\xi_n\uparrow f$) から，このとき f の積分は

$$\int_\Omega f d\mu=\lim_{n\to\infty}\int_\Omega \xi_n d\mu$$

で定められる．これは $\{\xi_n\}$ の選び方によらない．

なお，正値とは限らない可測函数 f は 2 つの正値可測函数 $f_+\equiv\max\{f,0\}$ と $f_-\equiv(-f)_+$ を使って $f=f_+-f_-$ と表せるから，この場合も，$\int_\Omega f_+ d\mu$ と $\int_\Omega f_- d\mu$ の一方が有限であれば

$$\int_\Omega f d\mu=\int_\Omega f_+ d\mu-\int_\Omega f_- d\mu$$

で f の積分が定められる．さらに，$\int_\Omega |f| d\mu=\int_\Omega f_+ d\mu+\int_\Omega f_- d\mu$ が有限値をとるとき，f は**可積分函数**であるという．f が複素数値函数の場合には，f を実数値函数 $\mathrm{Re}\, f$ と $\mathrm{Im}\, f$ に分割し，その各々に上の定義を適用すればよい．なお，測度空間 $(\Omega,\mathcal{F},\mu)$ における可積分函数の全体を $L^1(\Omega,\mathcal{F},\mu)$ あるいは単に $L^1(\mu)$ で表す：$L^1(\mu)\equiv\left\{f\in\mathfrak{M}_\mathcal{F};\ \int_\Omega |f| d\mu<\infty\right\}$.

以上の積分の定義より，積分に関する様々な命題は単函数に関して成立することを示せば十分であることが多い．ところで，$A\in\mathcal{F}$ における f の積分は $\int_A f d\mu\equiv\int_\Omega f\cdot 1_A d\mu$ で定める．

積分が定義できる $f,g\in\mathfrak{M}_\mathcal{F}$ に対して次が成立する：(1) $f\leqq g$ (μ-a.e.) $\Rightarrow\int_\Omega f d\mu$

$\leqq\int_\Omega g\,d\mu$. (2) $\int_\Omega(\alpha f+\beta g)d\mu=\alpha\int_\Omega f\,d\mu+\beta\int_\Omega g\,d\mu$ ($\forall\alpha,\beta\in \boldsymbol{C}$). (3) $|f|\leqq g$ (μ-a. e.) で $g\in L^1(\mu)\Rightarrow f\in L^1(\mu)$. (4) $f\in L^1(\mu)\Rightarrow|f|<\infty$($\mu$-a. e.).
(5) $\left|\int_\Omega f\,d\mu\right|\leqq\int_\Omega|f|d\mu$. (6) $\mu(A)=0\Rightarrow\int_A f\,d\mu=0$.

問 3 $f\in\mathfrak{M}_{\mathcal{F}}$, $f\geqq0$ に対して, $\int_\Omega f\,d\mu=\sup\left\{\int_\Omega\xi\,d\mu;\ 0\leqq\xi\leqq f, \xi\text{ は単函数}\right\}$ を示せ.

〈定理 3.3〉 単調収束定理：$\{f_n\}\subset\mathfrak{M}_{\mathcal{F}}$ が $0\leqq f_n\uparrow f$ (μ-a. e.) であれば

$$\lim_{n\to\infty}\int_\Omega f_n d\mu=\int_\Omega f\,d\mu.$$

略証 $\mu(N)=0$ なる N の元を除けば $f_n(\omega)\uparrow f(\omega)$ だから, Ω 上で $f_n\uparrow f$ と仮定してよい. $f_n\leqq f$ だから $\int_\Omega f_n d\mu\leqq\int_\Omega f\,d\mu$. よって, 逆向き $\lim_{n\to\infty}\int_\Omega f_n d\mu\geqq\int_\Omega f\,d\mu$ を示せばよい. 各 n に対して, 単函数の列 $\{\xi_{nk}\}$ が存在して, $0\leqq\xi_{nk}\uparrow f_n$ となるから, ここで, $\zeta_k=\max\{\xi_{nk};\ 1\leqq n\leqq k\}$ とおくと, $\zeta_k\uparrow f$ で

$$\int_\Omega f\,d\mu=\lim_{k\to\infty}\int_\Omega\zeta_k d\mu\leqq\lim_{k\to\infty}\int_\Omega f_k d\mu$$

がわかる. ▮

〈定理 3.4〉 Fatou の補題：$\{f_n\}\subset\mathfrak{M}_{\mathcal{F}}$ が $f_n\geqq0$ (μ-a. e.) であれば

$$\int_\Omega\underline{\lim}\ f_n d\mu\leqq\underline{\lim}\int_\Omega f_n d\mu.$$

証明 $\underline{\lim}\ f_n=\lim_{n\to\infty}\inf_{k\geqq n}f_k$ であるから $\inf_{k\geqq n}f_n\uparrow\underline{\lim}\ f_n$. よって〈定理 3.3〉より,

$$\int_\Omega\underline{\lim}\ f_n d\mu=\lim_{n\to\infty}\int_\Omega\inf_{k\geqq n}f_n d\mu.$$

また, $\inf_{k\geqq n}f_k\leqq f_n$ であるから,

$$\lim\int_\Omega\inf_{k\geqq n}f_k\leqq\underline{\lim}\int_\Omega f_n d\mu.\ ▮$$

〈定理 3.5〉 Lebesgue の有界収束定理：$\{f_n\}\subset\mathfrak{M}_{\mathcal{F}}$ がある $g\in L^1(\mu)$ に対して $|f_n|\leqq g$ (μ-a. e.) を満たし, $f_n\xrightarrow{\text{a. e.}}f$ であるとき

$$\int_\Omega f\,d\mu=\lim_{n\to\infty}\int_\Omega f_n d\mu.$$

略証 $|f_n|\leqq g$ より $|f|\leqq g$. ゆえに $g\pm f_n\geqq0$ に〈定理 3.4〉を適用すると

$$\int_\Omega(g\pm f)d\mu\leqq\underline{\lim}\int_\Omega(g\pm f_n)d\mu.$$

ここで, $\underline{\lim}\left(-\int_\Omega f_n d\mu\right)=-\overline{\lim}\int_\Omega f_n d\mu$ に注意すると命題が示される. ▮

問 4　$f_n \geqq 0$ (μ-a. e.) ならば $\sum_{n=1}^{\infty}\int_{\Omega} f_n d\mu = \int_{\Omega}\sum_{n=1}^{\infty} f_n d\mu$ を示せ．

問 5　$\{A_n\} \subset \mathcal{F}$ がたがいに素で，$A = \bigcup_{n=1}^{\infty} A_n$ かつ $f \in L^1(\mu)$ のとき，次を示せ．

$$\int_A f d\mu = \sum_{n=1}^{\infty}\int_{A_n} f d\mu.$$

3.5　直積空間と Fubini の定理

2 つの測度空間 $(X, \mathcal{F}, \mu)$ と $(Y, \mathcal{E}, \nu)$ において，$\mathcal{F}\otimes\mathcal{E}$ を集合 $\{A\times B;\ A\in\mathcal{F}, B\in\mathcal{E}\}$ より生成された σ-集合体とし，$E\in\mathcal{F}\otimes\mathcal{E}$ に対して，$E_x \equiv \{y\in Y;\ (x, y)\in E\}$，$E_y \equiv \{x\in X;\ (x, y)\in E\}$ と x-切片 y-切片を定めると

1°　μ, ν が σ-有限であれば，$E\in\mathcal{F}\otimes\mathcal{E}$ に対して次が成立する．

（1）　各 $x\in X$ に対して $E_x\in\mathcal{E}$ で $\nu(E_x)$ は $\mathcal{F}$-可測函数．

（2）　各 $y\in Y$ に対して $E_y\in\mathcal{F}$ で $\mu(E_y)$ は $\mathcal{E}$-可測函数．

（3）　$\int_X \nu(E_x) d\mu = \int_Y \mu(E_y) d\nu.$

この（3）の等式を $\mu\otimes\nu(E)$ とおくと，これは $\mathcal{F}\otimes\mathcal{E}$ 上の測度であり，μ と ν の直積測度とよぶ．積分の順序交換に関して次の結果は有用である．

註　直積測度の定義は，任意の $A\in\mathcal{F}$，$B\in\mathcal{E}$ に対して，$\varphi(A\times B) = \mu(A)\nu(B)$ を満たす $\mathcal{F}\otimes\mathcal{E}$ 上の測度として与えられる．

〈定理 3.6〉　Fubini の定理：$\mathcal{F}\otimes\mathcal{E}$-可測函数 f に対して，下記の（1）の 3 つの積分のうちひとつが有限ならば，他の 2 つも有限であり，（2）が成立する．

（1）　$\int_{X\times Y}|f| d\mu\otimes\nu,\ \int_X d\mu\int_Y d\nu|f|,\ \int_Y d\nu\int_X d\mu|f|.$

（2）　$\int_{X\times Y} f d\mu\otimes\nu = \int_X d\mu\int_Y d\nu f = \int_Y d\nu\int_X d\mu f.$

略証　f が正の可測函数のとき，1° より $\int_Y f d\nu$ が $\mathcal{F}$-可測で，$\int_X f d\mu$ が $\mathcal{E}$-可測であることがわかり，等式（2）が証明される．さらに，一般の $f\in\mathfrak{M}_{\mathcal{F}\otimes\mathcal{E}}$ では，$f\in L^1(\mu\otimes\nu)$ であれば $\int_Y f d\nu\in L^1(\mu)$，$\int_X f d\mu\in L^1(\nu)$ となるから，$f = f_+ - f_-$ と分解することにより（2）が示せる．∎

3.6　測度の分解と Radon-Nikodym の定理

μ を $(\Omega, \mathcal{F})$ 上の σ-加法的集合函数とし，$\{E_k;\ k=1, 2, \cdots, n\}$ を $E\in\mathcal{F}$ の有限分割 (i. e., $E_i\cap E_j = \emptyset$，$E = \bigcup_{k=1}^{n} E_k$) とする．このとき

$$|\mu|(E) \equiv \sup\left\{\sum_{k=1}^{n}|\mu(E_k)|;\ \{E_k\} \text{ は } E \text{ の有限分割}\right\}$$

$$\|\mu\| \equiv |\mu|(\Omega).$$

上の $\|\cdot\|$ を μ の全変動という.

1° μ が有限 (i. e., $|\mu(\Omega)|<\infty$) ならば, μ は有界であり (すなわち, ある定数 $M>0$ が存在して, $|\mu(A)|<M$, $\forall A\in\mathscr{F}$), かつ $|\mu|$ は有限測度になる.

2° **Jordan 分解**: $\mu_+(A)\equiv\frac{1}{2}\{|\mu|(A)+\mu(A)\}$, $\mu_-(A)\equiv\frac{1}{2}\{|\mu|(A)-\mu(A)\}$ ($\forall A\in\mathscr{F}$) の $\mu_\pm$ は $(\Omega,\mathscr{F})$ 上の σ-加法的集合函数で, $\mu=\mu_+-\mu_-$.

問 6 2° の μ_+ は $\mu_+(A)=\sup\{\mu(B);\ B\subset A,\ B\in\mathscr{F}\}$ に等しいことを示せ (なお, $\mu_-=(-\mu)_+$).

3° **Hahn 分解**: $(\Omega,\mathscr{F},\mu)$ において, Ω は μ に関する正集合 E (i. e., $\mu(A\cap E)\geqq 0$, $\forall A\in\mathscr{F}$) と負集合 F (i. e., $\mu(A\cap F)\leqq 0$, $\forall A\in\mathscr{F}$) に分解できる: $E\cap F=\emptyset$, $\Omega=E\cup F$. この分解は μ-零集合を除いて一意である.

略証 問6より $\mu_+(\Omega)=\sup\{\mu(E);\ E\subset\Omega\}$ であるから, 各 $n\in N$ に対して, $E_n\in\mathscr{F}$ が存在して

$$\mu_+(\Omega)\geqq\mu(E_n)\geqq\mu_+(\Omega)-\frac{1}{2^n}$$

となる. さらに, 任意の $G\subset E_n$ に対して, $\mu(G)\geqq-\frac{1}{2^n}$ が成立. また, $\Omega\supset\bigcup_{k=n}^{\infty}E_k$ であるから, 上の2つの不等式を使うと $\mu_+(\Omega)\geqq\mu(\bigcup_{k=n}^{\infty}E_k)\geqq\mu_+(\Omega)-\frac{1}{2^{n-1}}$ がわかる. $A_n=\bigcup_{k=n}^{\infty}E_k$ とおくと, $\{A_n\}$ は減少列. したがって, 単調クラス定理より $E\in\mathscr{F}$ が存在して $A_n\downarrow E$. この E に対して

$$\mu_+(\Omega)=\mu(E),\qquad \mu(A\cap E)\geqq 0\quad(\forall A\in\mathscr{F}).$$

また, $F=\Omega\backslash E$ とおくと,

$$\mu(A\cap F)\leqq 0\quad(\forall A\in\mathscr{F}).\ ▮$$

今, ν を $(\Omega,\mathscr{F})$ 上のもうひとつの σ-加法的集合函数とする. ある $A\in\mathscr{F}$ において, "$\mu(A)=0\Rightarrow\nu(A)=0$" が成立するとき, ν は μ に関して絶対連続であるといい, $\boldsymbol{\nu\ll\mu}$ と記す. また, Ω が $\Omega=E\cup F$, $E\cap F=\emptyset$ と分割できて, $\mu(E)=\nu(F)=0$ となるとき, μ と ν はたがいに特異であるといい, $\boldsymbol{\mu\perp\nu}$ と記す.

〈定理 3.7〉 **Radon-Nikodym の定理**:

(1) μ が σ-有限で ν が有限かつ $\nu\ll\mu$ であれば, $f\in L^1(\mu)$ が a. e. の意味で一意に存在して $\nu(A)=\int_A f\,d\mu$ ($\forall A\in\mathscr{F}$).

(2) μ と ν がともに σ-有限で $\nu\ll\mu$ であれば, 可測函数 f が a. e. の意味で一意に存在して $\nu(A)=\int_A f\,d\mu$ ($\forall A\in\mathscr{F}$).

略証 μ,ν とも有限で $\nu\geqq 0$ の場合に略証を与えよう. $\mathcal{G}=\Big\{g\in\mathfrak{M}_{\mathscr{F}};\ \int_E g\,d\mu\leqq\nu(E),$

$\forall E\in\mathcal{F}\}$ とおいて，$M=\sup\left\{\int_\Omega g\,d\mu;\ g\in\mathcal{G}\right\}$ とおくと，$f\in\mathcal{G}$ が存在して，$M=\int_\Omega f\,d\mu$ となることが示せる．ここで Hahn の分解定理を使うと，この f に対して

$$\nu(E)=\int_E f\,d\mu$$

となる．すなわち，上の f が求めるものである．▌

註 1 f を **Radon-Nikodym** (R. N. と略) **微分**といい，$\boldsymbol{f=d\nu/d\mu}$ で表す．

註 2 μ の σ-有限性は本質的である．

〈**定理 3.9**〉 **Lebesgue 分解**：σ-有限な μ,ν に対して，$\nu=\nu_{ac}+\nu_s$ と分解できる．ただし，$\nu_{ac}\ll\mu$，$\nu_s\,\delta\,\mu$ である．

さらに，次の結果も有用である．

4° **Vitali-Hahn-Saks の定理**：$(\Omega,\mathcal{F},\mu)$ を完備な σ-有限測度空間とし，$\{\nu_n\}$ を σ-加法的集合函数の列とする．このとき，$\nu_n\ll\mu$ で，$\{\nu_n(A)\}$ $(\forall A\in\mathcal{F})$ が収束すれば，$\{\nu_n\}$ は μ に関して一様絶対連続である．(すなわち，任意の $\varepsilon>0$ に対して，$\delta>0$ が存在して $\mu(A)<\delta$ ならば $|\nu_n(A)|<\varepsilon$ $(\forall n)$)．

5° **Nikodym の定理**：σ-加法的集合函数の列 $\{\mu_n\}$ において，$\{\mu_n(A)\}$ $(\forall A\in\mathcal{F})$ が収束すれば，$\mu(A)\equiv\lim_{n\to\infty}\mu_n(A)$ も σ-加法的集合函数となる．

3.7 測度論的確率論

可測空間 $(\Omega,\mathcal{F})$ 上の測度 μ で $\mu(\Omega)=1$ となるものを確率測度とよんだが，以下これを仮定する．可測函数 $f\colon\Omega\to\boldsymbol{R}$ (確率論ではこれを**確率変数**という) に対して，

$$F_f(x)\equiv\mu(\{\omega\in\Omega;\ f(\omega)\leq x\})\quad(=\mu(f\leq x)\ \text{と略})$$

を f の**分布函数**，任意の $Q\in\mathcal{B}_{\boldsymbol{R}}$ ($\boldsymbol{R}$上の Borel 集合の全体) に対して

$$\mu_f(Q)\equiv\mu(f^{-1}(Q))\quad(f^{-1}(Q)\equiv\{\omega\in\Omega;\ f(\omega)\in Q\})$$

をfに関する**確率分布という**．また，μ_f が Lebesgue 測度 dx に関して絶対連続であるとき，R. N. の定理より**密度函数** $p_f(x)$ が存在して

$$F_f(x)=\int_{-\infty}^{x}p_f(t)\,dt$$

とかける．なお，3つ組 $(\boldsymbol{R},\mathcal{B}_{\boldsymbol{R}},\mu_f)$ を f によって導かれた**確率空間**という．

元 $A\in\mathcal{F}$ を確率論では**事象**というが，事象 $\{A_\alpha;\alpha\in J\}$ および確率変数 $\{f_\alpha;\alpha\in J\}$ が**独立であるとは**，各々任意の $\alpha_1,\cdots,\alpha_n\in J$ と $Q_1,\cdots,Q_n\in\mathcal{B}_{\boldsymbol{R}}$ に対して

$$\mu\left(\bigcap_{k=1}^{n}A_{\alpha_k}\right)=\prod_{k=1}^{n}\mu(A_{\alpha_k}),$$

$$\mu\left(\bigcap_{k=1}^{n}f_k^{-1}(Q_k)\right)=\prod_{k=1}^{n}\mu(f_k^{-1}(Q_k))$$

を満たす場合である．

本節 3.3 項で確率変数の収束について論じたが，さらに，次の収束も重要である．確率変数の列 $\{f_n\}$ に対して

（1） **法則収束**：$f_n \xrightarrow{L} f \overset{\text{def}}{\Longleftrightarrow} F_f(x)$ の連続点において $F_{f_n}(x) \to F_f(x)$．

以下，確率論において基本的である概念および諸結果を列記しておこう．ただし，f, g, f_n は確率変数とする．

（2） **期待値**：$E(f) \equiv \int_\Omega f\,d\mu = \int_R x\,dF_f(x)$．

（3） **n 次積率**：$E(f^n) \equiv \int_\Omega f^n d\mu$．

（4） **分散**：$V(f) \equiv E(\{f-E(f)\}^2)$．

（5） **標準偏差**：$D(f) \equiv \sqrt{V(f)}$．

（6） **共分散**：$\mathrm{Cov}(f, g) \equiv E(\{f-E(f)\}\{g-E(g)\})$．

（7） **相関係数**：$\mathrm{Cor}(f, g) \equiv \mathrm{Cov}(f, g)/D(f)D(g)$．

1° **Chebyshev の不等式**：任意の $\varepsilon>0$ と $n \in N$ に対して，（1） $\mu(|f| \geqq \varepsilon) \leqq \dfrac{1}{\varepsilon^n} E(|f|^n)$；（2） $\mu(|f-E(f)| \geqq \varepsilon) \leqq \dfrac{1}{\varepsilon^2} V(f)$．

2° **Borel-Canteli の補題**：$\{A_n\} \subset \mathcal{F}$ に対して，（1） $\sum_{n=1}^{\infty} \mu(A_n) < \infty$ ならば $\mu(\overline{\lim}\, A_n)=0$；（2） $\{A_n\}$ が独立で，$\sum_{n=1}^{\infty} \mu(A_n)=\infty$ ならば $\mu(\overline{\lim}\, A_n)=1$．

3° **Kolmogorov の不等式**：$\{f_n\}$ が独立なとき，任意の $\varepsilon>0$ に対して，

$$\mu\left(\left\{\max_{1 \leq k \leq n}\left|\sum_{i=1}^{k}(f_i - E(f_i))\right| \geqq \varepsilon\right)\leqq \frac{1}{\varepsilon^2}\sum_{i=1}^{n} V(f_i).$$

4° **収束に関して**：（1） $f_n \xrightarrow{\mu} f \Rightarrow f_n \xrightarrow{L} f$．（2） $E(|f_n - f|^m) \to 0$ $(m \geqq 1) \Rightarrow f_n \xrightarrow{\mu} f$．

5° **中心極限定理**：$\{f_k\}$ が同じ分布に従う独立な確率変数で $V(f_k)=\sigma^2<\infty$ とする．$g_n = \sum_{k=1}^{n} f_k$，$h_n = \dfrac{g_n - E(g_n)}{\sqrt{n}\,\sigma}$ とおくと

$$F_{h_n}(x) \longrightarrow \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

6° **大数の弱法則**：$\{f_n\}$ が独立で，$E(f_k)<\infty$，$\dfrac{1}{n^2}\sum_{k=1}^{n} V(f_k) \to 0$ $(n \to \infty)$ を満たすとき，

$$\frac{1}{n}(g_n - E(g_n)) \xrightarrow{\mu} 0 \quad \left(g_n \equiv \sum_{k=1}^{n} f_k\right).$$

7° **大数の強法則**：$\{f_k\}$ が独立で，$E(f_k)<\infty$，$\sum_{n=1}^{\infty} \dfrac{V(f_n)}{n^2} < \infty$ を満たすとき，

$$\frac{1}{n}(g_n - E(g_n)) \xrightarrow{\text{a.e.}} 0 \quad \left(g_n \equiv \sum_{k=1}^{n} f_k\right).$$

4. Hilbert 空間とその作用素

この付章では Hilbert 空間上の種々な作用素の主要な性質について論じる.

4.1 Hilbert 空間

数体 $\boldsymbol{K}$($\boldsymbol{R}$ または $\boldsymbol{C}$) 上の線形空間 X の任意の元 x, y に対して, 以下の 3 つの条件を満たす数 $\langle x, y\rangle\in\boldsymbol{K}$ が定まるとき, $\langle x, y\rangle$ を x と y の内積といい, 内積をもつ線形空間を**内積空間**という. (i) $\langle x, x\rangle\geqq 0$, $=0\Longleftrightarrow x=0$; (ii) $\langle x, y\rangle=\overline{\langle y, x\rangle}$; (iii) $\langle x, \alpha y_1+y_2\rangle=\alpha\langle x, y_1\rangle+\langle x, y_2\rangle$, $\forall y_1, y_2\in X$, $\forall\alpha\in\boldsymbol{K}$.

〈定理 4.1〉 Schwarz の不等式：内積空間 X において, $|\langle x, y\rangle|\leqq\|x\|\cdot\|y\|$, $\forall x, y\in X$. 等号はある数 $\lambda\in\boldsymbol{K}$ が存在して, $y=\lambda x$ のとき成立, ただし, $\|x\|=\langle x, x\rangle^{1/2}$ である.

証明　$x=0$ のとき不等式の両辺$=0$ で明らか. $x\neq 0$ とする. $x'=x/\|x\|$ とおくと,

$$0\leqq\|\langle x', y\rangle x'-y\|^2=|\langle x', y\rangle|^2-2|\langle x', y\rangle|^2+\|y\|^2.$$

ゆえに $|\langle x', y\rangle|^2\leqq\|y\|^2$, したがって $|\langle x, y\rangle|\leqq\|x\|\cdot\|y\|$ を得る. この計算式から, 定理の不等式が等号ならば $y=\langle x', y\rangle x'=(\langle x', y\rangle/\|x\|)x$. 逆に $y=\lambda x$ ならば $|\langle x, y\rangle|=|\lambda|\cdot\|x\|^2=\|x\|\cdot\|y\|$. ▌

この不等式より, $\|\cdot\|$ は X 上にノルムを定めることが容易にわかる. つまり, (i) $\|x\|\geqq 0$, $=0\Longleftrightarrow x=0$, (ii) $\|\alpha x\|=|\alpha|\cdot\|x\|$, $\forall\alpha\in\boldsymbol{K}$, (iii) $\|x+y\|\leqq\|x\|+\|y\|$.

内積空間 X がノルム $\|\cdot\|$ に関して完備であるとき, X を **Hilbert 空間** (**$\boldsymbol{K}=\boldsymbol{R}$ なら実 Hilbert 空間, $\boldsymbol{K}=\boldsymbol{C}$ なら複素 Hilbert 空間**) といい, 以下, Hilbert 空間を $\mathscr{H}$ で表す.

ここで内積空間 X および Hilbert 空間 $\mathscr{H}$ における主要な性質を列記する.

1° **中線定理**：$\|x+y\|^2+\|x-y\|^2=2(\|x\|^2+\|y\|^2)$, $\forall x, y\in X$.

2° X が**ノルム空間** (i. e., ノルム $\|\cdot\|$ が定義されている空間) が 1° を満たせば, 任意の $x, y\in X$ に対して, $\langle x, y\rangle=\frac{1}{4}(\|x+y\|^2-\|x-y\|^2+i\|ix+y\|^2-i\|ix-y\|^2)$ (**偏極化**)によって, X に内積$\langle\cdot, \cdot\rangle$が導入される. 逆に, 内積空間 X はつねにこの等式を満たす. とくに X が **Banach 空間** (完備なノルム空間)で 1° を満たすとき, X は Hilbert 空間になる. ここで, Hilbert 空間 $\mathscr{H}$, および Banach 空間の例を挙げておこう.

例 1　$\mathscr{H}=\boldsymbol{R}^n$, $\langle x, y\rangle=\sum_{i=1}^{n}x_iy_i$; $\mathscr{H}=\boldsymbol{C}^n$, $\langle x, y\rangle=\sum_{i=1}^{n}\bar{x}_iy_i$.

例 2　数列空間 $\mathcal{S}=\{x=(x_n);\ x_n\in\boldsymbol{C}, n\in\boldsymbol{N}\}$ に対して, $l^2\subset\mathcal{S}$ を次のように定めると, これは Hilbert 空間になる.

$$l^2=\left\{x\in\mathcal{S};\ \langle x, x\rangle=\sum_{n=1}^{\infty}|x_n|^2<\infty\right\},\quad\langle x, y\rangle=\sum_{n=1}^{\infty}\bar{x}_ny_n.$$

例 3　$(\Omega, \mathscr{F}, \mu)$ を測度空間 (μ が σ-有限の場合も含む) とし, $\mathscr{F}$-可測函数の全体を $\mathfrak{M}$ とする. このとき,

$$L^2(\Omega, \mu)=\left\{x\in\mathfrak{M} \text{ の同値類}; \langle x, x\rangle=\int_\Omega |x|^2 d\mu<\infty\right\}$$

は内積 $\langle x, y\rangle=\int_\Omega \overline{x(\omega)}y(\omega)d\mu(\omega)$ に関して Hilbert 空間になる.

例 4 数列空間 $\mathcal{S}$ の次の部分空間 l^∞, c_0, l^p ($p\geqq1$) は Banach 空間である.

$$l^\infty=\{x\in\mathcal{S};\ \|x\|(\text{または } \|x\|_\infty)\equiv\sup_n|x_n|<\infty\},$$
$$c_0=\{x\in\mathcal{S};\ \lim_{n\to\infty} x_n=0\},$$
$$l^p=\left\{x\in\mathcal{S};\ \|x\|_p\equiv\left(\sum_{n=1}^\infty |x_n|^p\right)^{1/p}<\infty\right\},$$
$$f=\{x\in\mathcal{S};\ \text{有限個の } n \text{ を除き } x_n=0\}.$$

これらは $f\subset l^p\subset c_0\subset l^\infty$, $l^p\subset l^q$ $(1\leq p\leq q)$ の関係にある. また $f_0=\{x\in f; x_n\in \boldsymbol{Q}\}$ ($\boldsymbol{Q}$ は有理数の全体) とおくと, (i) f_0 の $\|\cdot\|_p$-閉包$=l^p$, (ii) f_0の$\|\cdot\|$-閉包$=c_0$, となるから (iii) l^p, c_0 は可分である. 一方, (iv) l^∞ は可分でない.

例 5 測度空間 $(\Omega, \mathcal{F}, \mu)$ に対して Banach 空間 $L^\infty(\Omega, \mu)$, $L^p(\Omega, \mu)(1\leq p<+\infty)$ を定めよう. $L^p(\Omega, \mu)$ は $\|x\|_p=\left(\int_\Omega |x|^p d\mu\right)^{1/p}<+\infty$ となる $x\in\mathfrak{M}$ (可測函数全体のなす線形空間) の同値類 (μ-a. e. による) であり, $L^\infty(\Omega, \mu)$ は $\|x\|$ (または $\|x\|_\infty$)$\equiv$ $\inf\{\alpha\in\overline{\boldsymbol{R}};\ |x(\omega)|\leq\alpha,\ \mu\text{-a. e.}\}<+\infty$ となる $x\in\mathfrak{M}$ の同値類である.

(1) $x\in L^\infty\cap L^1\Rightarrow x\in L^p$ $(p\geqq1)$.

(2) $\mu(\Omega)<\infty\Rightarrow L^\infty\subset L^q\subset L^p$ $(1\leqq p\leqq q)$, $\|x\|=\lim_{p\to\infty}\|x\|_p$ かつ $\overline{L^\infty}^{\|\cdot\|_p}=L^p$ $(p\geqq1)$.

(3) **Minkowski の不等式**: $\|x+y\|_p\leqq\|x\|_p+\|y\|_p$, $\forall x, y\in L^p$.

(4) **Hölder の不等式**: $\|xy\|_r\leqq\|x\|_p\|y\|_q$, $\forall x\in L^p$, $\forall y\in L^q$, $r^{-1}=p^{-1}+q^{-1}$.

Ω を局所コンパクト Hausdorff 空間とし, Ω 上の複素数値 (resp. 実数値) 連続函数の全体を $C(\Omega)$(resp. $C_R(\Omega)$) とかく. 今, $\|f\|=\sup\{|f(\omega)|; \omega\in\Omega\}$ とおくと, $C(\Omega)$ はノルム $\|\cdot\|$ に関して Banach 空間となる.

さて, Hilbert 空間の議論にもどろう. Hilbert 空間 $\mathcal{H}$ においてもとの可算列 $\{x_n\}_{n=1}^\infty(\subset\mathcal{H})$ に対して $x=\sum x_n$ とは, 任意の $\varepsilon>0$ に対して自然数 n_0 が存在して $\left\|x-\sum_{k=1}^n x_k\right\|<\varepsilon$ $(\forall n\geqq n_0)$ のときをいう. $\sum x_n$ が存在すれば一意である. これを $\{x_n\}$ の和という.

次に直交系の話に移ろう. 元 $x, y\in\mathcal{H}$ が $\langle x, y\rangle=0$ のとき x と y は直交するといい, $x\perp y$ で表す. $x\perp y\Longleftrightarrow y\perp x$ である. $\{x_n\}$ が $\langle x_i, x_j\rangle=0$ $(i\neq j)$ のとき**直交系**といい, さらに $\|x_n\|=1$ $(\forall n)$ のとき $\{x_n\}$ を**正規直交系** (**ONS** と略記) という. $\langle x_n, x\rangle=0$ $(\forall n)$ ならばつねに $x=0$ のとき $\{x_n\}$ を完全 (complete) といい, 完全正規直交系を **CONS** と略記する. Hilbert 空間の直交性と CONS に関する有用な結果について述べよう.

3° **Gram-Schmidt の直交化法**: $\{x_n\}\subset\mathcal{H}$ を一次独立な元の列とする. これに対して

$$\varphi_1 = x_1/\|x_1\|,$$
$$\varphi_2 = y_2/\|y_2\|, \qquad y_2 = x_2 - \langle\varphi_1, x_2\rangle\varphi_1,$$
$$\varphi_3 = y_3/\|y_3\|, \qquad y_3 = x_3 - \langle\varphi_1, x_3\rangle\varphi_1 - \langle x_3, \varphi_2\rangle\varphi_2$$
$$\cdots\cdots\cdots, \quad \cdots\cdots\cdots$$
$$\varphi_n = y_n/\|y_n\|, \quad y_n = x_n - \sum_{k=1}^{n-1}\langle\varphi_k, x_n\rangle\varphi_k, \qquad n=1,2,\cdots$$

とおくと $\{\varphi_n;\ n=1,2,\cdots\}$ は ONS である．

証明　数学的帰納法によって行う：$\|\varphi_n\|=1$ $(\forall n)$ は明白．

$$\langle\varphi_1, \varphi_2\rangle = \langle x_1, x_2 - \langle\varphi_1, x_2\rangle\varphi_1\rangle/\|x_1\|\cdot\|y_2\|$$
$$= \{\langle x_1, x_2\rangle - \langle x_1, \varphi_1\rangle\langle\varphi_1, x_2\rangle\}/\|x_1\|\cdot\|y_2\| = 0.$$

$\{\varphi_1, \varphi_2, \cdots, \varphi_n\}$ が ONS であるとする．このとき φ_{n+1} を加えて $\{\varphi_1, \varphi_2, \cdots, \varphi_{n+1}\}$ も ONS であることを示す．

$$\langle\varphi_k, \varphi_{n+1}\rangle = \frac{1}{\|y_{n+1}\|}\{\langle\varphi_k, x_{n+1}\rangle - \langle\varphi_k, \langle\varphi_k, x_{n+1}\rangle\varphi_k\rangle\} = 0 \quad (k \leq n). \ \blacksquare$$

4°　**Pythagorean.**　（1）　$x \perp y \Rightarrow \|x+y\|^2 = \|x\|^2 + \|y\|^2$.

（2）　有限な $\mathrm{ONS}\{\varphi_k\}_{k=1}^N$ に対して

$$\|x\|^2 = \left\|x - \sum_{k=1}^N \langle\varphi_k, x\rangle\varphi_k\right\|^2 + \sum_{k=1}^N |\langle\varphi_k, x\rangle|^2.$$

（2）の証明：$\sum = \sum_{k=1}^N$ とおくと，

$$\|x - \sum\langle\varphi_k, x\rangle\varphi_k\|^2 = \|x\|^2 + \sum\|\langle\varphi_k, x\rangle\varphi_k\|^2 - 2\mathrm{Re}\langle\sum\langle\varphi_k, x\rangle\varphi_k, x\rangle$$
$$= \|x\|^2 + \sum|\langle\varphi_k, x\rangle|^2 - 2\mathrm{Re}\sum|\langle\varphi_k, x\rangle|^2$$
$$= \|x\|^2 - \sum|\langle\varphi_k, x\rangle|^2. \ \blacksquare$$

5°　**Bessel の不等式**：ONS $\{\varphi_n\}$ に対し，$\sum|\langle\varphi_n, x\rangle|^2 \leqq \|x\|^2$（これは 4° の（2）から明らか）．

6°　$\mathrm{ONS}\{\varphi_n\} \subset \mathscr{H}$ に対して，次は同値である：

（1）　$\{\varphi_n\}$ は CONS (i. e.,$\langle\varphi_n, x\rangle = 0, \forall n \in \boldsymbol{N} \Rightarrow x = 0$).

（2）　**Fourier 展開**：$x = \sum\langle\varphi_n, x\rangle\varphi_n$，$\forall x \in \mathscr{H}$，ここで $\langle\varphi_n, x\rangle$ を x の **Fourier 係数**，右辺を **Fourier 級数**，この等式を $\mathrm{ONS}\{\varphi_n\}$ による x の **Fourier 展開**という．

（3）　**Riesz-Fischer の等式**：$\|x\|^2 = \sum|\langle\varphi_n, x\rangle|^2$.

（4）　**Parseval の等式**：$\langle x, y\rangle = \sum\langle x, \varphi_n\rangle\langle\varphi_n, y\rangle$.

証明　（1）⇨（2）：$y_n = \sum_{k=1}^n \langle\varphi_k, x\rangle\varphi_k$ とおくと，Bessel の不等式から $\|y_m - y_n\|^2 = \sum_{n+1}^m |\langle\varphi_k, x\rangle|^2 \to 0$ $(m \geq n \to \infty)$，ゆえに $y = \lim y_n = \sum\langle\varphi_k, x\rangle\varphi_k$ が存在する．また各 n に対して $\langle\varphi_n, y\rangle = \sum\langle\varphi_k, x\rangle\langle\varphi_n, \varphi_k\rangle = \langle\varphi_n, x\rangle$，すなわち，$\langle\varphi_n, x-y\rangle = 0$，$\forall n \in N$，ゆえに $x=y$．（2）⇨（3）：$\mathscr{H}$ のノルム収束の意味で $x = \lim_n \sum_{k=1}^n \langle\varphi_k, x\rangle\varphi_k$，ゆえに $\|x\|^2 = \lim_n \sum_{k=1}^n |\langle\varphi_k, x\rangle|^2 = \sum|\langle\varphi_n, x\rangle|^2$．（4）⟺（3）⟹（1）は明白である．$\blacksquare$

次に直交函数系 (ONS) の例を挙げよう．

例 6 （1） 三角函数系．$\mathscr{H}=L^2(-\pi,\pi)$ のとき，$\varphi_1(t)=\dfrac{1}{\sqrt{2\pi}}$，$\varphi_{2n}(t)=\dfrac{1}{\sqrt{\pi}}\cos nt$，$\varphi_{2n+1}=\dfrac{1}{\sqrt{\pi}}\sin nt$ のとき $\{\varphi_n;\ n=1,2,\cdots\}$ は CONS である．$L^2(-\pi,\pi)$ の函数の Fourier 展開は，通常この CONS が用いられる．

（2） 標本函数系．$\mathscr{H}=L^2(-\infty,\infty)$ のとき

$$\varphi_n(t)=\sqrt{\frac{W}{\pi}}\,\frac{\sin W(t-n\pi/W)}{W(t-n\pi/W)},\qquad n=0,\ \pm1,\ \pm2,\cdots$$

のとき $\{\varphi_n;\ n=0,\pm1,\pm2,\cdots\}$ は ONS であり，信号函数の標本化展開に用いられる．詳しくは第 6 章，6.2 節参照．

（3） $\mathscr{H}=L^2(-\pi,\pi)$ で $\left\{\dfrac{1}{\sqrt{2\pi}}\exp(int);\ n=0,\pm1,\pm,2\cdots\right\}$ は CONS.

（4） **Legendre の多項式：**

$$P_n(t)=\frac{1}{2^n n!}\,\frac{d^n}{dt^n}(t^2-1)^n\quad(n=0,1,2,\cdots)$$

に対して $\varphi_n(t)=\sqrt{(2n+1)/2}\,P_n(t)$ $(n=0,1,2,\cdots)$ は $\mathscr{H}=L^2(-1,1)$ で CONS.

（5） **Hermite 多項式**

$$H_n(t)=(-1)^n\exp(t^2)\frac{d^n}{dt^n}\exp(-t^2)$$

に対して $\varphi_n(t)=H_n(t)\exp(-t^2/2)/(2^n n!)^{1/2}\cdot\pi^{1/4}$，$n=0,1,2,\cdots$ は $\mathscr{H}=L^2(-\infty,\infty)$ で CONS.

（6） **Laguerre 多項式**

$$L_n(t)=e^t\frac{d^n}{dt^n}(t^ne^{-t})$$

に対して $\varphi_n(t)=L_n(t)e^{-t/2}/n!$，$n=1,2,\cdots$ は $\mathscr{H}=L^2(0,\infty)$ で CONS.

註 1 直交系の例は，ほかにいくらでもあるが，割合頻繁に用いられるものを列挙した．（4）の $\{\varphi_n(t)\}$ は $x_n(t)=t^n$，$n=0,1,2,\cdots$，の Gram-Schmidt の直交化法を用いても構成される．Hilbert 空間 l^2 では標準的な CONS は $\varphi_1=(1,0,0,\cdots)$，$\varphi_2=(0,1,0,\cdots)$，$\cdots$，$\varphi_n=(0,0,\cdots,1,0,\cdots)$，$\cdots$，ここで φ_n は n 番目が 1，他はすべて 0．

註 2 以上の議論では ONS$\{\varphi_n\}$ はすべて可算の範囲内で行った．これらのことはすべて非可算の $\{\varphi_j\}=\{\varphi_j;j\in J\}$ に対しても成立する．Hilbert 空間 $\mathscr{H}$ の元の集合 $\{x_j\}=\{x_j;j\in J\}$ に対して $x=\sum_J x_j$ とは，任意の $\varepsilon>0$ に対して，有限集合 $J_0\subset J$ が存在して $\|x-\sum_{J'}x_j\|<\varepsilon$ ($\forall$ 有限 J', $J_0\subset J'\subset J$) のときをいう．$\sum_J x_j$ は存在すれば一意である．これを可算の場合と同様 $\{x_j\}$ の和という．$\sum_J$ は x_j がスカラーの場合も同様に定義される．これは非可算の場合を含めた無限級数の収束の定義である．この場合，$\{j\in J;\ x_j\neq0\}$ は高々可算である．非可算$\{\varphi_j\}\subset\mathscr{H}$ が ONS であるとは $\langle\varphi_i,\varphi_j\rangle=\delta_{ij}$，CONS であるとは $\langle x,\varphi_j\rangle=0\ (\forall j)$ ならば $x=0$．このとき（1）Hilbert 空間 $\mathscr{H}$ ではつねに CONS が存在する．これは Zorn の補題を用いれば容易である．（2）$\mathscr{H}$ の

CONS が可算$\Longleftrightarrow\mathcal{H}$ は可分.

$\mathcal{H}$ の線形部分空間を単に**部分空間**といい，$\mathcal{H}$ のノルム $\|\cdot\|$ で閉じている部分空間を**閉部分空間**という．今，$\mathcal{K}$ を $\mathcal{H}$ の任意の部分集合とすると，$\mathcal{K}^{\perp}=\{x\in\mathcal{H};\langle x,y\rangle=0,\ \forall y\in\mathcal{K}\}$ は $\mathcal{H}$ の閉部分空間となり，これを $\mathcal{K}$ の**直交補空間**という．$\mathcal{H}$ の２つの閉部分空間 $\mathcal{K},\mathcal{L}$ において，$\langle x,y\rangle=0$ ($\forall x\in\mathcal{K},\forall y\in\mathcal{L}$) であるとき，$\mathcal{K}$ と $\mathcal{L}$ はたがいに直交するといい，$\mathcal{K}\perp\mathcal{L}$ で表す．$\mathcal{K}\perp\mathcal{L}\Longleftrightarrow\mathcal{L}\perp\mathcal{K}$ である．直和 $\mathcal{K}\oplus\mathcal{L}\equiv\{x+y;\ x\in\mathcal{K},y\in\mathcal{L}\}$ も $\mathcal{H}$ の閉部分空間となる．$\mathcal{H}$ の部分集合 $\mathcal{K}$ において

7°　$\mathcal{K}^{\perp\perp}=(\mathcal{K}^{\perp})^{\perp}$ は $\mathcal{K}$ を含む最小の閉部分空間である．

8°　ONS $\mathcal{K}=\{\varphi_j\}$ が完備 (CONS)$\Longleftrightarrow\mathcal{K}^{\perp}=\{0\}\Longleftrightarrow\mathcal{K}^{\perp\perp}=\mathcal{H}$．この 7°, 8° は次の 9° を用いて示される．

9°　**最短距離定理**：$\mathcal{K}$ が閉部分空間であるとき，$x\in\mathcal{H}$ に対して，$\delta=\inf\{\|x-y\|;\ y\in\mathcal{K}\}$ とおくと，$z\in\mathcal{K}$ が一意に存在して，$\delta=\|x-z\|$ となる．このとき $x\longmapsto z$ を $z=P_{\mathcal{K}}x$ とおき，$P_{\mathcal{K}}$ を $\mathcal{H}$ から $\mathcal{K}$ への**射影作用素**という．

証明　明らかに '$\delta=0\Longleftrightarrow x\in\mathcal{K}$ で $z=x$' となる．$\delta>0$ のとき $\{y_n\}\subset\mathcal{K}$ が存在して $\|x-y_n\|\to\delta$．したがって

$$\|y_m-y_n\|^2=\|(y_m-x)-(y_n-x)\|^2=2\|y_m-x\|^2+2\|y_n-x\|^2-\|y_m+y_n-2x\|^2$$

(1° 参照)

$$=2\|y_m-x\|^2+2\|y_n-x\|^2-4\left\|\frac{1}{2}(y_m+y_n)-x\right\|^2$$
$$\leqq 2\|y_n-x\|^2+2\|y_n-x\|^2-4\delta^2$$
$$\to 2\delta^2+2\delta^2-4\delta^2=0\quad(m,n\to\infty).$$

ゆえに $\{y_n\}$ は Cauchy 列，$\mathcal{H}$ は完備から $z\in\mathcal{H}$ が存在して $\|y_n-z\|\to 0$．また $\mathcal{K}$ は閉であるから $z\in\mathcal{K}$．さらに $\delta\leqq\|x-z\|=\lim\|x-y_n\|=\delta$，すなわち $\|x-z\|=\delta$．一意性は，もし $z'\in\mathcal{K}$ が $\|x-z'\|=\delta\Rightarrow 0\leq\|z-z'\|^2=\|(z-x)-(z'-x)\|^2=2\|z-x\|^2+2\|z'-x\|^2-4\left\|\frac{1}{2}(z+z')-x\right\|^2\leqq 0$，よって $z=z'$．▮

10°　部分集合 $\mathcal{K}\subset\mathcal{H}$ が，$x,y\in\mathcal{K}$，$0\leq\alpha\leq 1$ に対してつねに $\alpha x+(1-\alpha)y\in\mathcal{K}$ のとき $\mathcal{K}$ を**凸集合**という．閉凸集合 $\mathcal{K}\subset\mathcal{H}$ に対して 9° が証明を変更することなしに成立する．

〈**定理 4.2**〉　**射影定理**：$\mathcal{K}$ が $\mathcal{H}$ の閉部分空間であるとき，（１）$(x-P_{\mathcal{K}}x)\in\mathcal{K}^{\perp}$；（２）各 $x\in\mathcal{H}$ は一意に $x=y+z(y\in\mathcal{K},z\in\mathcal{K}^{\perp})$ と表され，$y=P_{\mathcal{K}}x,z=P_{\mathcal{K}^{\perp}}x$；（３）$\mathcal{H}=\mathcal{K}\oplus\mathcal{K}^{\perp}$（直和）が成立する．

証明　（１）：$u=x-P_{\mathcal{K}}x$ とおくと，9° の $\delta\geq 0$ を用い，任意の実数 t と $y\in\mathcal{K}$ に対し，$P_{\mathcal{K}}x+ty\in\mathcal{K}$ から $\delta^2\leq\|x-(P_{\mathcal{K}}+ty)\|^2=\|(x-P_{\mathcal{K}}x)-ty\|^2=\|u-ty\|^2=\|u\|^2-2t\,\mathrm{Re}\langle u,y\rangle+t^2\|y\|^2$，$\|u\|=\delta$ から $t^2\|y\|^2-2t\,\mathrm{Re}\langle u,y\rangle\geq 0$ ($\forall t\in\boldsymbol{R}$)．ゆえに $\mathrm{Re}\langle u,y\rangle=0$．この計算で t を it で置き換えると $\mathrm{Im}\langle u,y\rangle=0$．したがって $u\perp y$．（２），（３）は

(1)と Pythagorean から明白. ▌

11° 射影作用素 $P_{\mathcal{K}}$ は線形かつべき等 (i. e., $P_{\mathcal{K}}(\alpha x+\beta y)=\alpha P_{\mathcal{K}}+\beta P_{\mathcal{K}}y$, $P_{\mathcal{K}}{}^2=P_{\mathcal{K}}$).

12° 2つの Hilbert 空間 $\mathcal{H}_1$ と $\mathcal{H}_2$ が同型 ($\mathcal{H}_1\cong\mathcal{H}_2$ とかく) であるとは, 線形変換U: $\mathcal{H}_1\to\mathcal{H}_2$, $U\mathcal{H}_1=\mathcal{H}_2$, が存在して, $\langle Ux, Uy\rangle_{\mathcal{H}_2}=\langle x, y\rangle_{\mathcal{H}_1}$ ($\forall x, y\in\mathcal{H}_1$) となる場合をいう. とくに (1) $\dim\mathcal{H}=n<\infty$ ならば $\mathcal{H}\cong\boldsymbol{R}^n$, または $\mathcal{H}\cong\boldsymbol{C}^n$, (2) $\dim\mathcal{H}=\infty$ かつ $\mathcal{H}$ が可分$\Longleftrightarrow\mathcal{H}\cong l^2$. これらの証明は 6° を用いればよい.

13° Hilbert 空間 $\mathcal{H}$ において, 任意の $x\in\mathcal{H}$ に対して $f_x(y)=\langle x, y\rangle$, $y\in\mathcal{H}$ とおくと, $f_y\in\mathcal{H}^*$($\equiv\mathcal{H}$ 上の有界線形汎函数の全体, p. 162) であり, $\|f_y\|=\|y\|$. これは Schwarz の不等式と $f_y(y)=\langle y, y\rangle=\|y\|^2$ からいえる. この逆はとくに重要である:

〈定理 4.3〉 Riesz の表現定理: 各 $f\in\mathcal{H}^*$ に対して $x_f\in\mathcal{H}$ が一意に存在して, $f(y)=\langle x_f, y\rangle$, $\forall y\in\mathcal{H}$.

証明 $f\equiv 0$ ならば $x_f\equiv 0$ とすればよい. $f\neq 0$ とする. $N_f=\{y\in\mathcal{H};\ f(y)=0\}$ (f の零空間, 核ともいう) は閉部分空間である. $f\neq 0$ から $N_f\subsetneqq\mathcal{H}$. ゆえに〈定理 4.2〉より $0\neq z\in N_f{}^\perp$ が存在して $f(z)\neq 0$. 任意の $y\in\mathcal{H}$ は $y=y_0+\lambda z$ ($y_0\in N_f$) と一意に表される. これは次のようにしていえる:

$$y=\{y-(f(y)/f(z))z\}+(f(y)/f(z))z,$$

$$f(\text{右辺第一項})=f(y-(f(y)/f(z))z)=f(y)-f(y)=0.$$

ゆえに $y-(f(y)/f(z))z(\equiv y_0)\in N_f$. また$(f(y)/f(z))z(=\lambda z)\in N_f{}^\perp$. 分解 $y=y_0+\lambda z\in N_f\oplus N_f{}^\perp$ は一意である. この z を用いて, $x_f\equiv(\overline{f(z)}/\|z\|^2)z$ とおくと, $\forall y\in\mathcal{H}$ に対して

$$\begin{aligned}\langle x_f, y\rangle&=\langle x_f, y_0+\lambda z\rangle=\lambda\langle x_f, z\rangle=\lambda\langle(\overline{f(z)}/\|z\|^2)z, z\rangle\\&=\lambda f(z)=f(\lambda z)=f(y_0)+f(\lambda z)=f(y_0+\lambda z)=f(y).\end{aligned}$$

一意性: もし $f(x)=\langle x, y\rangle\forall y\in\mathcal{H}\Rightarrow\langle x, y\rangle=\langle x_f, y\rangle\forall y\Rightarrow x=x_f$. ▌

註 3 対応 $x\in\mathcal{H}\to f_x\in\mathcal{H}^*$ は共役線形等距離変換であり, $\mathcal{H}$ が実 Hilbert 空間ならば $\mathcal{H}\cong\mathcal{H}^*$ である. $x=f_x$ と同一視することにより $\mathcal{H}=\mathcal{H}^*$ と考えてよい.

註 4 Radon-Nikodym の定理 (付章3節の〈定理 3.7〉) は積分論において最も重要な定理のひとつである. これの証明はそれほど容易でないが, Riesz の表現定理を用いると以下のように容易である: μ, ν を有限測度で $\nu\ll\mu$ とし, $m=\mu+\nu$ とおく. 任意の $x\in B(\Omega)$ (Ω 上の有界 $\mathcal{F}$-可測函数全体のなす線形空間) に対して

$$|\mu(x)|=\left|\int_\Omega x\,d\mu\right|\leq\int_\Omega|x|\,d\mu\leq\int_\Omega|x|\,dm\leq m(\Omega)^{1/2}\left(\int_\Omega|x|^2dm\right)^{1/2}.$$

$\mu(\cdot)$ は $B(\Omega)$ 上の線形汎函数 (i. e., $\mu:\boldsymbol{B}(\Omega)\to\boldsymbol{C}$ で線形: $\mu(\alpha x+\beta y)=\alpha\mu(x)+\beta\mu(y)$, $\alpha, \beta\in\boldsymbol{C}$) であり, また, $x, y\in B(\Omega)$ が “$x=y\overset{\text{def}}{\Longleftrightarrow}x(\cdot)=y(\cdot)$ m-a. e.” とすると $B(\Omega)$ は L^2 で稠密. したがって, $\mu(\cdot)$ は L^2 上の有界線形汎函数 (同じく $\mu(\cdot)$)

で表す）に一意に拡張される：$|\mu(x)|\leq\|x\|_2$. ゆえに Riesz の定理により一意に $x_\mu\in L^2$ が存在して $\mu(y)=\langle x_\mu, y\rangle$, $\forall y\in L^2$. 同様に $x_\nu\in L^2$ が存在し $\nu(y)=\langle x_\nu, y\rangle$, $\forall y\in L^2$. ゆえに $\langle 1, y\rangle=m(y)=\mu(y)+\nu(y)=\langle x_\mu, y\rangle+\langle x_\nu, y\rangle=\langle x_\mu+x_\nu, y\rangle$ から $x_\mu+x_\nu=1$ (m-a. e.). 一方 $\langle x_\mu, y\rangle=0$ $(0\leq y\in L^2)\Rightarrow\int y\,d\mu=0\Rightarrow y=0$ (μ-a. e.). したがって $x_\mu{}^{-1}(\omega)=1/x_\mu(\omega)$ が μ-a. e. で存在し，$x_0=x_\mu{}^{-1}\cdot x_\nu\in L^1(\Omega, \mu)$ が示され，目的の $x_0=d\nu/d\mu$ を得る．▮

この項の最後に，2つの Hilbert 空間 $\mathscr{H}_1$ と $\mathscr{H}_2$ のテンソル積について論じよう．各 $x_1\in\mathscr{H}_1$, $x_2\in\mathscr{H}_2$ に対して，直積空間 $\mathscr{H}_1\times\mathscr{H}_2=\{(y_1, y_2);\ y_1\in\mathscr{H}_1, \in y_2\mathscr{H}_2\}$ 上の共役双線形写像 $x_1\otimes x_2$ を

$$x_1\otimes x_2(y_1, y_2)=\langle y_1, x_1\rangle_{\mathscr{H}_1}\langle y_2, x_2\rangle_{\mathscr{H}_2}$$

で定義し，$\mathscr{E}$ をこの共役双線形写像の有限個の線形結合全体のなす線形空間とし，$\mathscr{E}$ 上に $\langle x_1\otimes x_2, y_1\otimes y_2\rangle=\langle x_1, y_1\rangle_{\mathscr{H}_1}\langle x_2, y_2\rangle_{\mathscr{H}_2}$ を定めると，これは $\mathscr{E}$ 上の内積を与える．この内積で $\mathscr{E}$ を完備化したものを $\mathscr{H}_1$ と $\mathscr{H}_2$ の**テンソル積 Hilbert 空間**といい，$\mathscr{H}_1\otimes\mathscr{H}_2$ で表す．$\{\varphi_j\}\{\psi_k\}$ をそれぞれ $\mathscr{H}_1, \mathscr{H}_2$ の CONS とすると，$\{\varphi_j\otimes\psi_k\}$ が $\mathscr{H}_1\otimes\mathscr{H}_2$ の CONS になることが証明される．次のテンソル積は重要である．今 $(X, \mu), (Y, \nu)$ を2つの測度空間，$\mathscr{H}$ を可分な Hilbert 空間とすると，（1）$L^2(X, \mu)\otimes L^2(Y, \nu)\cong L^2(X\times Y, \mu\otimes\nu)$；（2）$L^2(X, \mu)\otimes\mathscr{H}\cong L^2(X, \mu;\ \mathscr{H})\left(=\left\{\sum_{k=1}^N f_k x_k;\ f_k\in L^2, x_k\in\mathscr{H}\right\}\right.$ の $\langle\cdot|\cdot\rangle$ による閉包；ただし $\left.\langle f_k x_k|g_l x_l\rangle=\int_X \bar{f}_k g_l\,d\mu\langle x_k, x_l\rangle\right)$.

4.2 非有界作用素

$\mathscr{H}, \mathscr{H}_1, \mathscr{H}_2$ を Hilbert 空間，A, B 等はその上の作用素であるとする．$\mathscr{H}_1$ の部分空間 $\mathscr{D}$ 上で定義され $\mathscr{H}_2$ に値をとる写像 A が．$A(\alpha x+\beta y)=\alpha Ax+\beta Ay$ ($\forall x, y\in\mathscr{D}$, $\forall\alpha, \beta\in\boldsymbol{K}$) を満たすとき，$A$ を $\mathscr{H}_1$ から $\mathscr{H}_2$ への**線形作用素**，または単に**作用素**という．$\mathscr{D}$ を A の**定義域**といい，$\mathbf{dom}\,\boldsymbol{A}$ で表し，また集合 $\mathbf{ran}\,\boldsymbol{A}=\{Ax; x\in\mathrm{dom}\,A\}$ を A の**値域**という．さらに，$\mathscr{G}(A)=\{x\oplus Ax;\ x\in\mathrm{dom}\,A\}$ を A の**グラフ**といい，この $\mathscr{G}(A)$ が直和空間 $\mathscr{H}_1\oplus\mathscr{H}_2$ で閉集合のとき，A を**閉作用素**という．これは $\|x_n\to x\|\to 0$, $\|Ax_n-y\|\to 0$ ならば $x\in\mathrm{dom}\,A$ かつ $y=Ax$ となることと同値である．2つの作用素 A, B において，$A\subset B$ (A は B に含まれる)$\Longleftrightarrow\mathscr{G}(A)\subset\mathscr{G}(B)$ で定義する．A を包む閉作用素 B が存在するとき，A を**閉化可能な作用素（可閉作用素）**といい，そのような B のうちで最小のものを A の**閉包**といい，$\bar{A}$ で表す．

1° 次の（1），（2）が成立：（1）$A\subset B\Longleftrightarrow\mathrm{dom}\,A\subset\mathrm{dom}\,B$ かつ $\mathrm{dom}\,A$ では $A=B$; （2）A が可閉$\Longleftrightarrow\mathscr{G}(\bar{A})=\overline{\mathscr{G}(A)}$.

A を $\mathscr{H}$ **から** $\mathscr{H}$ **への**（これを単に $\mathscr{H}$ **上の**という）作用素とする．“$x\in\mathscr{H}$ に対して $x^*\in\mathscr{H}$ が存在して等式 $\langle x, Ay\rangle=\langle x^*, y\rangle$, $\forall y\in\mathrm{dom}\,A$” という命題が成り立つような $x\in\mathscr{H}$ の全体を $\mathscr{D}$ とおく．このとき，$\overline{\mathrm{dom}\,A}=\mathscr{H}$ であれば（このような A を**稠密**

定義域をもつ作用素という)，上記命題における対応 $x\in\mathscr{D}\to x^*\in\mathscr{H}$ は一意である．ここで，$A^*x\equiv x^*$, $x\in\mathscr{D}$ とおくと，A^* は $\mathscr{H}$ 上の作用素となる．この A^* を A の**共役作用素**といい，$A\subset A^*$ である A を**対称作用素**または **Hermit 作用素**という．とくに，$\mathrm{dom}\,A=\mathrm{dom}\,A^*$ であるとき，$A=A^*$ とかき，A は**自己共役作用素**という．

2° $\mathrm{dom}\,A=\mathscr{H}$ のとき次の（1）～（3）が成立する：（1）A が対称作用素ならば可閉；（2）作用素 A が可閉ならば A^{**} が存在して，$\bar{A}=A^{**}$（対称作用素 A の A^{**} が自己共役のとき，A を**本質的に自己共役**であるという）；（3）A が閉作用素ならば A^*A が存在し，これが自己共役である．

4.3 有界作用素

Hilbert 空間 $\mathscr{H}$ 上の作用素 A が $\mathrm{dom}\,A=\mathscr{H}$ かつ実数 $M>0$ が存在して $\|Ax\|\leqq M\|x\|$, $\forall x\in\mathscr{H}$ であるとき，A を**有界作用素**という．$\mathscr{H}$ 上の有界作用素の全体を $\boldsymbol{B}(\mathscr{H})$ で表し，また $\mathscr{H}_1$ から $\mathscr{H}_2$ への有界作用素の全体は $\boldsymbol{B}(\mathscr{H}_1,\mathscr{H}_2)$ で表す．なお，$\|A\|\equiv\sup\{\|Ax\|;\|x\|=1\}$ で作用素 A のノルム $\|\cdot\|$ が定まり，$\boldsymbol{B}(\mathscr{H})$ における線形演算は各点ごとの和 $((\alpha A+\beta B)x=\alpha Ax+\beta By)$ で定義して，$\boldsymbol{B}(\mathscr{H})$ は Banach 空間となる．

以下，A,B を $\boldsymbol{B}(\mathscr{H})$ の元とする．作用素 A が，$\langle x,Ax\rangle\geqq 0$ $(\forall x\in\mathscr{H})$ のとき，A を**正の作用素**といい，$A\geqq 0$ で表す．このとき，正の作用素 B が一意に存在して $A=B^2$ とかけ，この B は $A^{1/2},\sqrt{A}$ などと表される．なお，複素 Hilbert 空間上の正の作用素はつねに自己共役作用素であるが，実 Hilbert 空間上では正の作用素でも自己共役であるとは限らない．

次の作用素は重要である．（1）A が**正規作用素** $\overset{\text{def}}{\Longleftrightarrow} A^*A=AA^*$．（2）$A$ が**射影作用素** $\overset{\text{def}}{\Longleftrightarrow} A=A^*=A^2$．（3）$A$ が**等距離作用素** $\overset{\text{def}}{\Longleftrightarrow}\|Ax\|=\|x\|$ $(\forall x\in\mathscr{H})$．（4）$A$ が**ユニタリー作用素** $\overset{\text{def}}{\Longleftrightarrow}\langle Ax,Ay\rangle=\langle x,y\rangle$ $(\forall x,y\in\mathscr{H})$ かつ $\mathrm{ran}\,A=\mathscr{H}$．（5）$A$ が**半等距離作用素** $\overset{\text{def}}{\Longleftrightarrow} A$ が $(\ker A)^\perp$ 上で等距離作用素（$(\ker A)^\perp$ を A の**初期空間**，$\mathrm{ran}\,A$ を A の**終空間**といい，$A^*A=P_{(\ker A)^\perp}$, $AA^*=P_{\mathrm{ran}\,A}$．ここで $P_{\mathscr{L}}$ は $\mathscr{H}$ からその部分空間 $\mathscr{L}$ への射影作用素である）．

〈定理 4.4〉 **極分解定理**：$A\in\boldsymbol{B}(\mathscr{H})$ に対して，半等距離作用素 W が存在して；（i）$A=W|A|$;（ii）$|A|=W^*A$;（iii）$A^*=W^*|A^*|$;（iv）$|A^*|=W|A|W^*$．ここで，$|A|=\sqrt{A^*A}$ である．なお，この分解は一意で，W: $|A|x\longrightarrow Ax$ の初期空間は $\overline{\mathrm{ran}\,|A|}$ で終空間は $\overline{\mathrm{ran}\,A}$ である．さらに，A が正規作用素であるか，$\mathscr{H}$ が有限次元であるときは，W はユニタリー作用素になる．

註 A が非有界作用素でも，A が $\overline{\mathrm{dom}\,A}=\mathscr{H}$ を満たす閉作用素であれば上と同様の定理が成立する．

4.4 スペクトル

$\mathcal{H}$上の作用素 A; $\mathrm{dom}\,A\to\mathrm{ran}\,A$ が逆元 A^{-1}: $\mathrm{ran}\,A\to\mathrm{dom}\,A$ をもつのは A が1対1写像の場合であるが，ある複素数 λ に対して，$(A-\lambda I)^{-1}$ の存在に関連して，複素数の全体 $\boldsymbol{C}$ が以下のように分類される．なお，$\mathrm{ran}(A-\lambda I)=\mathrm{dom}(A-\lambda I)^{-1}$ である．

（1） **リゾルベントの集合**：$Re(A)=\{\lambda\in\boldsymbol{C};\ (A-\lambda I)^{-1}\in\boldsymbol{B}(\mathcal{H})\}$.

（2） **スペクトルの集合**：$Sp(A)=\boldsymbol{C}-Re(A)$（$Re(A)$ の補集合）.

スペクトル集合 $Sp(A)$ は，さらに以下の3つの集合に分けられる．

（3） **点スペクトルの集合**：$Sp^{(p)}(A)=\{\lambda\in\boldsymbol{C};\ (A-\lambda I)^{-1}$ が存在しない$\}$.

（4） **連続スペクトルの集合**：$Sp^{(c)}(A)=\{\lambda\in\boldsymbol{C};\ (A-\lambda I)^{-1}$ は存在するが非有界で $\overline{\mathrm{ran}(A-\lambda I)}=\mathcal{H}\}$.

（5） **剰余スペクトルの集合**：$Sp^{(r)}(A)=\{\lambda\in\boldsymbol{C};\ (A-\lambda I)^{-1}$ が存在するが非有界で $\overline{\mathrm{ran}(A-\lambda I)}\neq\mathcal{H}\}$.

各スペクトル集合はたがいに素であり，$Sp(A)=Sp^{(p)}(A)\cup Sp^{(c)}(A)\cup Sp^{(r)}(A)$.

ある $\lambda\in\boldsymbol{C}$ において，点列 $\{x_n\}\subset\mathcal{H}(\|x_n\|\geqq 1)$ が存在して，$\|(A-\lambda I)x_n\|\to 0$ となるとき，λ を**近似点スペクトル**という．このスペクトルの集合を $Sp^{(\mathrm{a.p.})}(A)$ で表す．このとき，（1） $Sp^{(p)}(A)\cup Sp^{(c)}(A)\subset Sp^{(\mathrm{a.p.})}(A)\subset Sp(A)$，（2） $Sp^{(r)}(A)=\emptyset$ であれば，$Sp^{(\mathrm{a.p.})}(A)=Sp^{(p)}(A)$ が成立する．

$Sp^{(p)}(A)$ の元 λ を A の**固有値**といい，この λ に対して，$Ax=\lambda x$ となる $x\in\mathcal{H}$ を固有値 λ に関する**固有ベクトル**という．固有ベクトル x の全体のつくる $\mathcal{H}$ の部分空間 $\mathcal{H}_\lambda$ を λ の**固有空間**といい，その次元 $\dim\mathcal{H}_\lambda$ を λ の**重複度**（または**多重度**）という．重複度が1の固有値を**単一**（または**非縮退**）スペクトルとよぶ．

$\mathcal{H}$ 上の射影作用素の族 $\{E_\lambda; -\infty\leq\lambda\leq\infty\}$ が（i） $E_\lambda E_\mu=E_\mu E_\lambda=E_\lambda\ (\lambda\leq\mu)$，（ii） $E_\lambda\to 0\ (\lambda\to-\infty)$，(iii) $E_\lambda\to I\ (\lambda\to+\infty)$ を満たすとき，$\{E_\lambda\}$ を**スペクトル測度**または**単位の分解**という．以下 $\{E_\lambda\}$ はスペクトル測度とする．$\{E_\lambda\}$ は次の性質を有する．

1° 右強連続性 $E_{\lambda_n}\downarrow E_\lambda\ (\lambda_n\downarrow\lambda)$ をもち，左強極限 $E_{\lambda-0}$ が存在する：$E_{\lambda_n}\uparrow E_{\lambda-0}(\lambda_n\uparrow\lambda))$.

2° $\left\{x\in\mathcal{H};\ \int_{\boldsymbol{R}}\lambda d\|E_\lambda x\|^2<+\infty\right\}^{-}=\mathcal{H}$.

3° $A=\int_{\boldsymbol{R}}\lambda dE_\lambda$ とおく $\left(\text{i. e., }\langle x, Ay\rangle=\int_{\boldsymbol{R}}\lambda d\langle x, E_\lambda, y\rangle,\ \forall x, y\in\mathcal{H}\right)$と，$A$ は自己共役作用素となる．

3° の逆である次の定理が**スペクトル分解定理**といわれるものである．

〈定理 4.5〉 自己共役作用素 A に対して，スペクトル測度 $\{E_\lambda\}$ が一意に存在して，$A=\int_{\boldsymbol{R}}\lambda dE_\lambda$ と表せる．

この定理において，"$E_{\lambda-0}\neq E_\lambda\Longleftrightarrow\lambda\in Sp^{(p)}(A)$" であり，このとき，$\mathrm{ran}(E_\lambda-E_{\lambda-0})$

は λ の固有空間になる．また，上の作用素 A において，$Sp(A)=Sp^{(p)}(A)=\{\lambda_n\}$ のとき，$P_n=E_{\lambda_n}-E_{\lambda_n-0}$ とおくと，$P_nP_m=\delta_{nm}P_n$で，$A=\sum_n \lambda_n P_n$ と表せる．

4.5 $\boldsymbol{F}(\mathcal{H})$ および $\boldsymbol{C}(\mathcal{H})$

Hilbert 空間 $\mathcal{H}$ の任意の元 x, y に対して，作用素 $x\otimes\bar{y}$ を

$$(x\otimes\bar{y})z=\langle y, z\rangle x, \qquad \forall z\in\mathcal{H}$$

で定めると，$x\otimes\bar{y}\in\boldsymbol{B}(\mathcal{H})$，かつ，$\operatorname{ran} x\otimes\bar{y}=\{\lambda x; \lambda\in\boldsymbol{C}\}$ は $x\neq 0$ のとき 1 次元閉部分空間であり，$x=0$ のときは一点集合 $\{0\}$ である．$x\otimes\bar{y}$ の作用素ノルム $\|x\otimes\bar{y}\|$ は $\|x\|\|y\|$ に等しい．さらに，次の性質が容易に証明される．

1° $(x\otimes\bar{y})^*=y\otimes\bar{x}$，$(\alpha x)\otimes(\overline{\beta y})=\alpha\bar{\beta}x\otimes\bar{y}$，$\forall\alpha, \beta\in\boldsymbol{C}$．

2° $(x+y)\otimes\bar{z}=x\otimes\bar{z}+y\otimes\bar{z}$，$x\otimes\overline{(y+z)}=x\otimes\bar{y}+x\otimes\bar{z}$．

3° $(x_1\otimes\bar{y}_1)(x_2\otimes\bar{y}_2)=\langle y_1, x_2\rangle x_1\otimes\bar{y}_2$．

4° $A(x\otimes\bar{y})=Ax\otimes\bar{y}$，$(x\otimes\bar{y})A=x\otimes\overline{A^*y}$，$\forall A\in\boldsymbol{B}(\mathcal{H})$．

この作用素 $x\otimes\bar{y}$ を x と y の **Schatten 積**（あるいは von Neumann–Schatten 積）という．これは **Dirac の記号：ブラ・ケット**を使うと，$x\otimes\bar{y}=|x\rangle\langle y|$ と表せる．この Dirac 記号では，内積は $\langle x, y\rangle=\langle x|y\rangle$ とかかれることが多い．

$\{x_j; j\in J\}$，$\{y_j; j\in J\}$を $\mathcal{H}$ の 2 つの ONS とし，$\mathcal{J}=\{\alpha;$ 有限集合 $\alpha\subset J\}$とすると，$A_\alpha=\sum_{j\in\alpha}\lambda_j x_j\otimes\bar{y}_j$ $(\sup_{j\in J}|\lambda_j|\leqq M<\infty)$ はある $A\in\boldsymbol{B}(\mathcal{H})$ に一意に強収束し(i. e., $\|A_\alpha x-Ax\|\to 0$)，この A を $A=\sum_{j\in J}\lambda_j x_j\otimes\bar{y}_j$ で表す．この A に関して，次が成立する．

5. $A^*=\sum_j\bar{\lambda}_j y_j\otimes\bar{x}_j$，$A^*A=\sum_j|\lambda_j|^2 y_j\otimes\bar{y}_j$，$AA^*=\sum_j|\lambda_j|^2 x_j\otimes\bar{x}_j$，および $|A|=\sum_j|\lambda_j|y_j\otimes\bar{y}_j$．

6° $\|A\|=\|A^*\|=\|A^*A\|^{1/2}=\||A|\|$．

7° $A^{-1}\in\boldsymbol{B}(\mathcal{H})\Longleftrightarrow\{x_j\}, \{y_j\}$ は，$\mathcal{H}$ は CONS で，定数 $c_1\geqq c_2>0$ が存在して，$c_1\geqq|\lambda_i|\geqq c_2$ $(\forall j)$ のとき，$A^{-1}=\sum_j\lambda_j^{-1}\bar{y}_j\otimes x_j$

〈**定理 4.6**〉 $\mathcal{H}$ 上の作用素 A において，

（1） A が射影作用素$\Longleftrightarrow\overline{\operatorname{ran} A}$ の CONS$\{x_j\}$ に対して，$A=\sum_j x_j\otimes\bar{x}_j$．

（2） A がユニタリー作用素$\Longleftrightarrow\mathcal{H}$ の CONS $\{x_j\}, \{y_j\}$ が存在して，$A=\sum x_j\otimes\bar{y}_j$．

（3） A が等距離作用素$\Longleftrightarrow\mathcal{H}$ の ONS$\{x_j\}$ と CONS$\{y_j\}$が存在して，$A=\sum x_j\otimes\bar{y}_j$．

（4） A が半等距離作用素$\Longleftrightarrow\mathcal{H}$ の ONS$\{x_j\}, \{y_j\}$ が存在して，$A=\sum x_j\otimes\bar{y}_j$．

ところで，$\dim(\operatorname{ran} A)<\infty$ である作用素 $A\in\boldsymbol{B}(\mathcal{H})$ を**有限階作用素**といい，この作用素の全体を $\boldsymbol{F}(\mathcal{H})$ で表す．$A\in\boldsymbol{F}(\mathcal{H})$ であることと，$n\in\boldsymbol{N}$, $\{\lambda_j\}\subset\boldsymbol{R}$, ONS$\{x_j\}$ および $\{y_i\}$ が存在して $A=\sum_{j=1}^n\lambda_j x_j\otimes\bar{y}_j$ と表せることは同値である．

2 つの Hilbert 空間 $\mathcal{H}_1, \mathcal{H}_2$ の間の作用素 A: $\mathcal{H}_1\to\mathcal{H}_2$ が**完全連続**（以下 C.C. と略）または**コンパクト**であるとは，有界集合 $\mathcal{L}\subset\mathcal{H}$ の A による像 $A\mathcal{L}$ が全有界の場合をいう．C.C. 作用素の全体を $\boldsymbol{C}(\mathcal{H}_1, \mathcal{H}_2)$ で表す．

8° A が C. C. $\Longleftrightarrow x_n \xrightarrow{w} x$ (i. e., $\langle y, x_n \rangle \to \langle y, x \rangle$, $\forall y \in \mathcal{H}$. この収束を弱 (weak) 収束という) ならば $Ax_n \xrightarrow{s} Ax$ (i. e., $\|Ax_n - Ax\| \to 0$; 強 (strong) 収束).

9° $\{A_n\} \subset \boldsymbol{C}(\mathcal{H}_1, \mathcal{H}_2)$ で $A_n \xrightarrow{u} A$ (i. e., 作用素 (一様) ノルム $\|Q\| \equiv \sup\{\|Qx\|;\ x \in \mathcal{H}_1,\ \|x\|=1\}$ において, $\|A_n - A\| \to 0$; 一様収束またはノルム収束) で $A \in \boldsymbol{B}(\mathcal{H}_1, \mathcal{H}_2)$ ($\|A\| < \infty$ なる作用素の全体) であるとき, $A \in \boldsymbol{C}(\mathcal{H}_1, \mathcal{H}_2)$.

10° $\boldsymbol{C}(\mathcal{H}_1, \mathcal{H}_2)$ は作用素ノルムに関して, Banach 空間 $\boldsymbol{B}(\mathcal{H}_1, \mathcal{H}_2)$ の閉部分空間であり, したがって $\boldsymbol{C}(\mathcal{H}_1, \mathcal{H}_2)$ 自身 Banach 空間である.

以下, $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$ の場合, すなわち, $\boldsymbol{C}(\mathcal{H}_1, \mathcal{H}_2) = \boldsymbol{C}(\mathcal{H}, \mathcal{H}) \equiv \boldsymbol{C}(\mathcal{H})$ の元に関して議論しよう. 上記の 8°～10° に加えて,

11° $\boldsymbol{C}(\mathcal{H})$ は $\boldsymbol{B}(\mathcal{H})$ の両側イデアルである.

12° $A \in \boldsymbol{B}(\mathcal{H})$ において, $A, A^*, |A|, |A^*|$ のいずれかひとつが C. C. ならば, 他も C. C. である.

13° $\lambda_j \to 0$ の点列 $\{\lambda_j\} \subset \boldsymbol{C}$ と 2 つの ONS$\{x_j\}$, $\{y_j\}$ に対して, $A = \sum \lambda_j x_j \otimes \bar{y}_j$ は C. C. である.

14° $A \in \boldsymbol{C}(\mathcal{H})$ ならば, $x \in \mathcal{H}$ が存在して, $|\langle x, Ax \rangle| = \max\{|\langle y, Ay \rangle|;\ y \in \mathcal{H},\ \|y\| = 1\}$. さらに $A \in \boldsymbol{C}(\mathcal{H})$ が自己共役ならば, $x \in \mathcal{H}$ が存在して, $\|A\| = |\langle x, Ax \rangle|$. また, この x は A の固有ベクトルである. すなわち, $Ax = \lambda x$ (λ は $\|A\|$ または $-\|A\|$).

〈定理 4.7〉 固有値展開定理：自己共役な $A \in \boldsymbol{C}(\mathcal{H})$ は次の性質を有する.

(1) $Sp(A) = Sp^{(p)}(A) \cup \{0\} \subset \boldsymbol{R}$ で集合 $Sp(A)$ は高々可算である. また, 0 以外の固有値は有限な重複度をもち, 次のように整列化できる. λ_1 を $|\lambda_1| = \|A\|$ ととり, その重複度を n_1 とすると, $\lambda_1 = \lambda_2 = \cdots = \lambda_{n_1}$, 次に λ_{n_1+1} を $|\lambda_{n_1+1}| = \max\{|\lambda|;\lambda \in Sp(A) - \{\lambda_1\}\}$ ととり, その重複度を n_2 とすると, $\lambda_{n_2+1} = \lambda_{n_1+2} = \cdots = \lambda_{n_1+n_2}$, 以下帰納的に λ_n を定める. このとき, 列 $\{\lambda_1, \lambda_2, \cdots\}$ は有限回で終わるか, $|\lambda_n| \downarrow 0$ である.

(2) 各 $\lambda_n \in Sp^{(p)}(A)$ における固有ベクトル $x_n (\|x_n\| = 1)$ を $x_n \perp x_m (m \neq n)$ と選んで, $A = \sum \lambda_n x_n \otimes \bar{x}_n$ と表せる. 固有値に縮退がなければ, 上記の分解は一意である.

(3) $\mathcal{L} = \overline{\mathrm{span}}\{x_n;\ n = 1, 2, \cdots\}$, $\mathcal{K} = \{z \in \mathcal{H};\ Az = 0\}$ とおくと, $\mathcal{L} = \mathcal{K}^\perp$ であり,

$$y = \sum \langle x_n, y \rangle + P_{\mathcal{K}} y, \qquad Ay = \sum \lambda_n \langle x_n, y \rangle x_n \quad (\forall y \in \mathcal{H}).$$

〈定理 4.8〉 min-max の原理：$A \in \boldsymbol{C}(\mathcal{H})$ が正のとき, A の固有値 $\{\lambda_n\}$ は次の min-max 形式で求められる.

$$\lambda_n = \min\{\max\{\langle Ax, x \rangle;\ \|x\| = 1,\ x \perp \mathcal{L}\};\ \dim \mathcal{L} = n-1\}$$

ここで, $\mathcal{L}$ は $(n-1)$ 次元の $\mathcal{H}$ の閉部分空間である (λ_1 は 14° で求まる λ と同じ固有値である).

15° $A, B \in \boldsymbol{C}(\mathcal{H})$ ともに自己共役とする. $AB = BA$ であることと, ONS$\{x_n\} \subset \mathcal{H}$

が存在して $A=\sum\lambda_n x_n\otimes\bar{x}_n$, $B=\sum\mu_n x_n\otimes\bar{x}_n$ と表せることは同値である.

16° すべての $A\in \boldsymbol{C}(\mathscr{H})$ に対して一意に $0\leqq\lambda_n\downarrow 0$ なる点列 $\{\lambda_n\}$ が存在して（これは $|A|$ の固有値），適当に ONS$\{x_n\}$, $\{y_n\}$ を選ぶと，$A=\sum_n\lambda_n x_n\otimes\bar{y}_n$ と表せる．逆に，上式で表される A は 13° で述べたようにつねに C. C. である．

17° 自己共役作用素 A において，ある自然数 n に対して，$A^n\in\boldsymbol{C}(\mathscr{H})$ であれば，A それ自身 C. C. である．

なお，これまでの議論より，$\boldsymbol{F}(\mathscr{H})\subset\boldsymbol{C}(\mathscr{H})$ であることは明らかであろう．

4.6 $S(\mathscr{H})$

$A\in\boldsymbol{B}(\mathscr{H})$ と，任意の2つの CONS$\{x_j\}$, $\{y_j\}\subset\mathscr{H}$ に対して，次の3つの無限級数

$$\sum_j\|Ax_j\|^2,\qquad \sum_j\|A^*y_j\|^2,\qquad \sum_{i,j}|\langle y_j, Ax_i\rangle|^2$$

のいずれかひとつが収束すれば他も収束し，極限値は一致する．このとき，A を**Schmidt 作用素**（または **Hilbert-Schmidt 作用素**）といい，Schmidt 作用素の全体を $S(\mathscr{H})$ で表す．

任意の CONS$\{x_j\}$ と $A\in S(\mathscr{H})$ に対して，$\|A\|_2=(\sum\|Ax_j\|^2)^{1/2}$ とおくと，$\|\cdot\|_2$ はノルム (**Schmidt ノルム**)となり，これに関して $S(\mathscr{H})$ は Banach 空間になる．$S(\mathscr{H})$ の元 A, B は次の性質を有する．

1° $\|x\otimes\bar{y}\|_2=\|x\|\|y\|$, $\forall x, y\in\mathscr{H}$.

2° $\|A\|\leqq\|A\|_2$, $\|A^*\|_2=\|A\|_2$, $\|AB\|_2\leqq\|A\|\|B\|_2$, $\|BA\|_2\leqq\|A\|\|B\|_2$. よって，$S(\mathscr{H})$ は $\boldsymbol{B}(\mathscr{H})$ の両側イデアルである．

3° $A\in S(\mathscr{H})\Rightarrow|A|\in S(\mathscr{H})$, $\||A|\|_2=\|A\|_2$.

なお，任意の $A, B\in S(\mathscr{H})$ と，任意の CONS$\{x_j\}$ に対して，$\langle\!\langle A, B\rangle\!\rangle=\sum\langle Ax_j, Bx_j\rangle$ とおくと，これは $\{x_j\}$ の選び方によらず，絶対収束し，$S(\mathscr{H})$ はこの内積 $\langle\!\langle\cdot,\cdot\rangle\!\rangle$ に関して，Hilbert 空間になる．さらに，$\langle\!\langle\cdot,\cdot\rangle\!\rangle$ は次の性質を有する．

4° $\langle\!\langle x\otimes\bar{y}, u\otimes\bar{v}\rangle\!\rangle=\langle x, u\rangle\langle v, y\rangle$, $\forall x, y, u, v\in\mathscr{H}$.

5° $A, B\in S(\mathscr{H})$. $C\in\boldsymbol{B}(\mathscr{H})$ に対して，$\langle\!\langle A^*, B^*\rangle\!\rangle=\langle\!\langle B, A\rangle\!\rangle$, $\langle\!\langle CA, B\rangle\!\rangle=\langle\!\langle A, C^*B\rangle\!\rangle$, $\langle\!\langle AC, B\rangle\!\rangle=\langle\!\langle A, BC^*\rangle\!\rangle$.

なお，次の結果は重要である．

6° $S(\mathscr{H})\subset\boldsymbol{C}(\mathscr{H})$.

したがって，任意の $A\in S(\mathscr{H})$ は $A=\sum_j\lambda_j x_j\otimes\bar{y}_j$ と表され，また，C. C. 作用素 $A=\sum_j\lambda_j x_j\otimes\bar{y}_j$ が $A\in S(\mathscr{H})$ である必要十分条件は $\sum_j|\lambda_j|^2<\infty$ であることもわかる．

4.7 $\boldsymbol{T}(\mathscr{H})$

任意の CONS$\{x_j\}$ に対して，無限級数 $\sum_j\langle x_j, Ax_j\rangle$ が絶対収束するとき，その値は CONS$\{x_j\}$ のとり方によらない．この $A\in\boldsymbol{B}(\mathscr{H})$ を**トレースクラスの作用素**といい，こ

れの全体を $T(\mathcal{H})$ で表す．以下，$\sum_j \langle x_j, Ax_j \rangle$ を $\mathrm{tr}\,A$ とかく．

1° 次は同値である．（1）$A \in \boldsymbol{T}(\mathcal{H})$．（2）$|A| \in \boldsymbol{T}(\mathcal{H})$．（3）$|A|^{1/2} \in S(\mathcal{H})$．（4）$\mathrm{tr}\,|A| < \infty$．

2° 任意の $A, B \in \boldsymbol{T}(\mathcal{H})$ と任意の $\alpha, \beta \in \boldsymbol{C}$ に対して $\alpha A + \beta B \in \boldsymbol{T}(\mathcal{H})$，$A^* \in \boldsymbol{T}(\mathcal{H})$．$\|A\|_1 = \mathrm{tr}\,|A| \, (\forall A \in \boldsymbol{T}(\mathcal{H}))$ とおくと，この $\|\cdot\|_1$ はノルムになり，$\boldsymbol{T}(\mathcal{H})$ は Banach 空間となる．

3° $\mathrm{tr}\, x \otimes \bar{y} = \langle y, x \rangle$，$\|x \otimes \bar{y}\|_1 = \|x\| \|y\|$，$\forall x, y \in \mathcal{H}$．

4° $\boldsymbol{T}(\mathcal{H})$ は $\boldsymbol{B}(\mathcal{H})$ の両側イデアルである．

5° $\|A\| \leqq \|A\|_2 \leqq \|A\|_1$．

6° $A \in \boldsymbol{T}(\mathcal{H}) \Longleftrightarrow B,\ C \in \boldsymbol{S}(\mathcal{H})$ が存在して，$A = BC$．

7° 任意の $A = \sum \lambda_j x_j \otimes y_j \in \boldsymbol{C}(\mathcal{H})$，$(\lambda_j \geqq 0)$ が $\boldsymbol{T}(\mathcal{H})$ の元 $\Longleftrightarrow \sum \lambda_j < \infty$．このとき，$\|A\|_1 = \sum \lambda_j$．

4°, 5° より $\boldsymbol{T}(\mathcal{H}) \subset \boldsymbol{S}(\mathcal{H})$ であることがわかる．

この項の最後に，空間 $\boldsymbol{B}(\mathcal{H})$，$\boldsymbol{C}(\mathcal{H})$，$\boldsymbol{F}(\mathcal{H})$，$S(\mathcal{H})$，$\boldsymbol{T}(\mathcal{H})$ の関連について論じよう．これらは＊代数（第2巻の付章2節2.1項参照）であり，さらに，

〈定理 4.9〉 $\boldsymbol{F}(\mathcal{H}) \subset \boldsymbol{T}(\mathcal{H}) \subset \boldsymbol{S}(\mathcal{H}) \subset \boldsymbol{C}(\mathcal{H}) \subset \boldsymbol{B}(\mathcal{H})$，の包含関係がある（ここで包含関係⊂は部分代数，両側イデアルという代数的意味合いも含まれている）．さらに，

（1） $\boldsymbol{F}(\mathcal{H})$ は他の各々の空間の最小イデアルである．また $\boldsymbol{C}(\mathcal{H})$ は $\boldsymbol{B}(\mathcal{H})$ の一様ノルムで閉じたイデアルであり，とくに $\mathcal{H}$ が可分ならば，最大イデアルである．

（2） 次が成立する．

（i）$\overline{\boldsymbol{F}(\mathcal{H})}^{\|\cdot\|_1} = \boldsymbol{T}(\mathcal{H})$．（ii）$\overline{\boldsymbol{F}(\mathcal{H})}^{\|\cdot\|_2} = S(\mathcal{H})$．（iii）$\overline{\boldsymbol{F}(\mathcal{H})}^{\|\cdot\|} = \boldsymbol{C}(\mathcal{H})$．

（3） 各 $\boldsymbol{F}(\mathcal{H})$，$\boldsymbol{T}(\mathcal{H})$，$S(\mathcal{H})$，$\boldsymbol{C}(\mathcal{H})$ は $\boldsymbol{B}(\mathcal{H})$ で弱位相（収束が $A_n \xrightarrow{w} A$ で得られる位相（第2巻の付章2節2.6項））に関して稠密．

今，Banach 空間 X 上の有界線形汎函数 φ（すなわち φ: $X \to \boldsymbol{C}$ が線形かつ $|\varphi(A)| \leqq M\|A\|$, $0 < M < +\infty$）の集合を X^* で表す．このとき任意の $\varphi \in \boldsymbol{C}(\mathcal{H})^*$ に対して，一意に $T_\varphi \in \boldsymbol{T}(\mathcal{H})$ が存在して，$\varphi(A) = \mathrm{tr}\, T_\varphi A \, (\forall A \in \boldsymbol{C}(\mathcal{H}))$ とかける．逆に，任意の $T \in \boldsymbol{T}(\mathcal{H})$ に対して，$\varphi_T(A) = \mathrm{tr}\, TA \, (\forall A \in \boldsymbol{C}(\mathcal{H}))$ とおくと，$\varphi_T \in \boldsymbol{C}(\mathcal{H})^*$ である．また，$\|T\|_1 = \|\varphi_T\| \equiv \sup\{|\varphi_T(A)|;\ A \in \boldsymbol{C}(\mathcal{H}),\ \|A\| = 1\}$ が満たされる．同様に，任意の $\psi \in \boldsymbol{T}(\mathcal{H})^*$ に対して一意に $B_\psi \in \boldsymbol{B}(\mathcal{H})$ が存在して，$\psi(A) = \mathrm{tr}\, AB_\psi$ $(\forall A \in \boldsymbol{T}(\mathcal{H}))$．逆に，任意の $B \in \boldsymbol{B}(\mathcal{H})$ に対して，$\psi_B(A) = \mathrm{tr}\, AB$ $(\forall A \in \boldsymbol{T}(\mathcal{H}))$ とおくと，$\psi_B \in \boldsymbol{T}(\mathcal{H})^*$ となり，$\|B\| = \|\psi_B\|$．

$T \in \boldsymbol{T}(\mathcal{H})$ と $\varphi_T \in \boldsymbol{C}(\mathcal{H})^*$，$B \in \boldsymbol{B}(\mathcal{H})$ と $\psi_B \in \boldsymbol{T}(\mathcal{H})^*$ を同一視することにより，次の定理が得られる．

〈定理 4.10〉 次が成立．（1）$\boldsymbol{C}(\mathcal{H})^* = \boldsymbol{T}(\mathcal{H})$．（2）$\boldsymbol{T}(\mathcal{H})^* = \boldsymbol{B}(\mathcal{H})$．（3）$S(\mathcal{H})^*$

$=\boldsymbol{S}(\mathcal{H})$. （4） $\boldsymbol{B}(\mathcal{H})^*=\boldsymbol{T}(\mathcal{H})\oplus\boldsymbol{C}(\mathcal{H})^{\perp}$. ただし， $\boldsymbol{C}(\mathcal{H})^{\perp}$ は集合 $\{\varphi\in\boldsymbol{B}(\mathcal{H})^*;\ \varphi(A)=0,\ \forall A\in\boldsymbol{C}(\mathcal{H})\}$ である.

5. 線形位相空間における基本定理

位相空間 X が $\boldsymbol{K}$ 上の線形空間であり，写像 $(x,y)\in X\times X\longrightarrow x+y\in X$，および $(\alpha,x)\in\boldsymbol{K}\times X\to\alpha x\in X$ が連続であるとき，X を**線形位相空間**という．以下，線形位相空間の基本的性質をごく簡単にまとめておく.

線形空間 X 上の写像 p: $X\to\boldsymbol{K}^+$ が，任意の $x,y\in X$ および $\lambda\in\boldsymbol{K}$ に対して，（i） $p(x+y)\leqq p(x)+p(y)$，（ii） $p(\lambda x)=|\lambda|p(x)$ を満たすとき，p を X 上の**半ノルム**という．半ノルムの族 $\{p_j; j\in J\}$ が X の元を**分離する**とは，すべての $j\in J$ に対して，$p_j(x)=0$ ならば $x=0$ となることである．また X 上に位相を，"$x_\alpha\to x \overset{\text{def}}{\Longleftrightarrow} p_j(x_\alpha-x)\to0$ $(\forall j\in J)$" で導入すると，X は**局所凸** (i. e., $0\in X$ の近傍基の要素が凸) な線形位相空間になる．逆に，局所凸な線形位相空間にはこのような半ノルムの族が存在する．さらに，X の元を分離する半ノルムの族が存在することと，Xが**局所凸 Hausdorff 空間**であることは同値である．したがって，ノルム空間は局所凸 Hausdorff 空間である.

〈定理 5.1〉 **Hahn-Banach 拡張定理**：Xを実 (resp. 複素) 線形空間，p を X 上の半ノルム，Y を X の部分空間，φを任意 $y\in Y$ に対して，$\varphi(y)\leqq p(y)$ (resp. $|\varphi(y)|\leqq p(y)$) を満たす Y 上の実 (resp. 複素) 線形汎函数とする．このとき，X 上の実 (resp. 複素) 線形汎函数 $\tilde{\varphi}$ が存在して，$\tilde{\varphi}\upharpoonright Y$($Y$ への制限)$=\varphi$ かつ $\tilde{\varphi}(x)\leqq p(x)$ (resp. $|\tilde{\varphi}(x)|\leqq p(x)$), $\forall x\in X$.

この定理の系として，次の結果は有用である．ただし，以下 X をノルム空間，Y を X の部分空間，X^* を X 上の有界線形汎函数の全体とする.

1° Y 上の有界線形汎函数 φ_0 に対して，$\varphi\in X^*$ が存在し，$\varphi_0=\varphi\upharpoonright Y$ かつ $\|\varphi_0\|_Y=\|\varphi\|$ が成り立つ.

2° $x\in X$, $\|x\|\neq0$ に対して，$\varphi(x)=\|\varphi\|\cdot\|x\|$ となる $\varphi\in X^*$ が存在する.

3° $x\in X$ と Y の距離を r とするとき，$\|\varphi\|\leqq1$, $\varphi(x)=r$, $\varphi(y)=0$ $(\forall y\in Y)$ を満たす $\varphi\in X^*$ が存在する.

〈定理 5.2〉 **Banach-Steinhaus** (一様有界性) 定理：X を Banach 空間，Y をノルム空間とし，$\mathcal{F}$ を $\boldsymbol{B}(X,Y)$ (X から Y への有界線形変換の全体) の部分集合とする．このとき，任意の $x\in X$ で $\{\|Tx\|_Y;\ T\in\mathcal{F}\}$ が有界であれば，$\mathcal{F}$ は一様有界である (i. e., $\sup\{\|T\|;\ T\in\mathcal{F}\}<+\infty$).

以下，しばらく X, Y を Banach 空間とする.

4° ξ: $X\times Y\to\boldsymbol{C}$ において，$\xi(x,\cdot)\in Y^*$, $\xi(\cdot,y)\in X^*$ のとき，$\xi(\cdot,\cdot)$ は同時連続である (i. e., $x_n\to0$, $y_n\to0\Rightarrow\xi(x_n,y_n)\to0$).

5° **開写象定理**：$T\in \boldsymbol{B}(X, Y)$ が上への写像のとき，T は開写像である（すなわち，T は開集合を開集合へ移す）.

6° **閉グラフ定理**：$T\in \boldsymbol{B}(X, Y)$ であることと，T のグラフ $\mathcal{G}=\{(x, Tx);\ x\in X\}$ が閉集合であることとは同値.

ここで，弱位相について簡単に記しておこう．ある集合 X とある位相空間 Y において，$\mathcal{F}$ を X から Y への写像の族とするとき，$\mathcal{F}$ によって導入された X 上の**弱位相**（または **$\mathcal{F}$-弱位相**）とは，各 $\varphi\in\mathcal{F}$ を連続にする最も弱い位相のことである．この位相は $\{\varphi^{-1}(O);\ \varphi\in\mathcal{F}, O$ は Y の開集合$\}$ によって生成され，この位相におけるネット $\{x_\alpha;\alpha\in J\}$ の収束（弱収束といい，$x_\alpha \xrightarrow{w} x$ とかく）は $\lim_\alpha \varphi(x_\alpha)=\varphi(x)$ $(\forall\varphi\in\mathcal{F})$ によって決められている．したがって，弱位相は**各点収束の位相**である．なお，Y が Hausdorff 空間であり，$\mathcal{F}$ が X の元を**分離する**（i. e., $\varphi(x)=\varphi(y)$, $\forall\varphi\in\mathcal{F}\Rightarrow x=y$）ならば，$X$の弱位相も Hausdorff 位相である.

さて，$\mathcal{F}\subset \boldsymbol{L}(X, \boldsymbol{C})$（$X$ から $\boldsymbol{C}$ への線形変換の全体）で $\mathcal{F}$ が X の元を分離すれば，X 上の $\mathcal{F}$-弱位相（$\sigma(X, \mathcal{F})$ 位相とかく）は $\mathcal{F}$ の各元を連続にする X 上の最も弱い位相である．この $\sigma(X, \mathcal{F})$ 位相は半ノルム $\{p_\varphi;\ p_\varphi(x)=|\varphi(x)|, \varphi\in\mathcal{F}\}$ によって生成される局所凸 Hausdorff 位相である．同様に，$\sigma(X, X^*)$ 位相，$\sigma(X^*, X)$ 位相も定義され，それぞれ X の弱位相，X^* の**弱＊位相**という．X_1 を Banach 空間Xの単位球（i. e., $X_1=\{x\in X;\ \|x\|\leqq 1\}$）とすると

〈定理 5.3〉 **Banach-Alaoglu 定理**：$(X^*)_1$ は弱＊位相でコンパクトである.

局所コンパクト Hausdorff 空間 X において，$C_0(X)$ の部分集合 $\mathcal{A}$ が部分代数(i. e., $f, g\in\mathcal{A}\Rightarrow fg\in\mathcal{A}$）であるとする．また集合 $K(\mathcal{A})$ を $\{x\in X;\ f(x)=0,\ f\in\mathcal{A}\}$ と定める．このとき

〈定理 5.4〉 **Stone-Weierstrass 定理**：$\mathcal{A}\subset C_0(X)$ が X の元を分離し，$K(\mathcal{A})=\emptyset$ のとき，$\overline{\mathcal{A}}^{\|\cdot\|}=C_0(X)$．また，$X$ がコンパクトの場合は，$\mathcal{A}\subset C(X)$ が X の元を分離し，$\mathcal{A}\ni 1$ であれば $\overline{\mathcal{A}}^{\|\cdot\|}=C(X)$ となる.

線形空間 X の凸集合 K の点 x が，いかなる K の2点 x_1, x_2 と数 $\lambda\in(0,1)$ に対しても，$x=\lambda x_1+(1-\lambda)x_2$ と表せない場合，xを K の**端点**（extreme point）であるといい，端点の集合を $\mathbf{ex}\,\boldsymbol{K}$ で表す．また，X のある部分集合 Y の凸包（convex hull）とは Y を含む最小の凸集合のことであり，これを $\mathbf{co}\,\boldsymbol{Y}$ で表し，その閉包を $\overline{\mathrm{co}}\,Y$ で表す．このとき次の有用な定理が成立する.

〈定理 5.5〉 **Krein-Milman 定理**：局所凸線形位相空間 X のコンパクトな凸集合 K において，(1) $\mathrm{ex}\,K\neq\emptyset$;（2）$K=\overline{\mathrm{co}}\,\mathrm{ex}\,K$.

その他の重要な定理を列記しておこう．以下，とくに断らなければ X を Banach 空

間，Y をその部分集合とする．

7° **Grothendieck 定理**：$\varphi\in X^{**}$ が弱位相で連続であることと $\varphi\restriction X^*{}_1$ が相対弱$*$位相で連続であることは同値．

8° **Eberlein-Šmulian 定理**：Y が相対弱コンパクト（すなわち Y の弱位相での閉包 $\bar{Y}^w$ がコンパクト）であることと Y が弱点列コンパクト（すなわち，任意の $\{y_n\}\subset Y$ に対して，部分列 $\{y_{n_k}\}$ と $y\in Y$ が存在して，$y_{n_k}\overset{w}{\to}y$）であることは同値．

9° **Krein-Šmulian 定理**：Kを X^* の凸部分集合とする．このとき，Kが弱$*$位相で閉であることと $K\cap rX^*{}_1$ $(r>0)$ が弱$*$位相で閉であることは同値 $(rX^*{}_1\equiv\{\varphi\in X^*;\ \|\varphi\|\leqq r\})$.

10° **Kakutani-Markov の不動点定理**：K をある線形位相空間 X の空でないコンパクトな凸部分集合とする．さらに，ξ を X から X への連続なアフィン写像で，$\xi(K)\subset K$ を満たすものとする．このとき，$\xi(x_0)=x_0$ となる $x_0\in K$ が存在する．

最後に Bochner 積分について概説する．$(\Omega, \mathcal{F}, \mu)$ をある測度空間，X を Banch 空間，f を Ω から X への写像とする．このとき，

（1） f が（X-値）**単函数**であるとは，$\mu(A_k)<\infty$ となる有限個のたがいに素な可測集合 $\{A_k\}_{k=1}^n$ に対して，$f(\omega)=x_k$ $(x_k\in X, \neq 0)$, $\forall\omega\in A_k$, かつ $f(\omega)=0$, $\forall\omega\in\Omega-\bigcup_{k=1}^n A_k$ となる場合をいう．

（2） f が**強可測**であるとは，単函数の列 $\{f_n\}$ が存在して，$\|f_n(\omega)-f(\omega)\|\to 0$ (μ-a. e.) となる場合をいう．

（3） f が**弱可測**であるとは，任意の $\varphi\in X^*$ に対して，$\omega\longrightarrow\varphi(f(\omega))$ が可測函数となる場合をいう．

当然，“（2）⇨（3）”が成立する．なお，この可測性に関して，

11° X が可分なとき，$f\colon\Omega\to X$ が強可測であることと弱可測であることは同値．

さて，Bochner 積分の定義に移ろう．$f\colon\Omega\to X$ を単函数とし，$f(\omega)=x_k$ $(\forall\omega\in A_k)$ とする．このとき，f の Ω 上の積分を

$$\int_\Omega f\,d\mu\equiv\sum_{k=1}^n x_k\mu(A_k)$$

と定め，これを f の Bochner 積分という．さらに，一般の函数 $f\colon\Omega\to X$ に対して，単函数の列 $\{f_n\}$ が存在して，$\|f_n(\omega)-f(\omega)\|\to 0$ (μ-a. e.) および $\int_\Omega\|f_n-f\|d\mu\to 0$ が満たされるとき，f は **Bochner 可積分**であるといい，積分

$$\int_\Omega f\,d\mu=s\text{-}\lim_{n\to\infty}\int_\Omega f_n\,d\mu$$

を **Bochner 積分**という．

12° 強可測な函数 $f\colon\Omega\to X$ が Bochner 可積分であることと $\omega\longmapsto\|f(\omega)\|$ が可積分函数であることとは同値．

この章と前章の結果は文献［R. 1, S. 3, Y. 2］などに詳しく論じられている．

参　考　文　献

ここでは，一般的な参考書と本書で引用した論文，および本書の内容と関連の深いもののみを挙げる．

[A.1] L. M. Abramov: On the entropy of a flow (Russian), Dokl. Akad. Nauk. SSSR, 128 (1959), 873-875.

[A.2] J. Aczel, B, B. Forte and C. T. Ng: Why the Shannon and Hartely entropys are "natural"? Adv. Appl. Prob. 6 (1974), 131-146.

[A.3] R. L. Adler: Ergodic and mixing properties of infinite memory channels, Proc. Amer. Math. Soc. 12 (1961), 924-930.

[A.4] R. L. Adler: A note on the entropy of skew product transformations, Proc. Amer. Math. Soc. 14 (1963), 665-669.

[A.5] R. L. Adler, A. Konheim and M. McAndrew: Topological entropy, Trans. Amer. Math. Soc. 114 (1965), 309-319.

[A.6] R. L. Adler and B. Marcus: Topological entropy and equivalence of dynamical systems, Memoirs Amer. Math. Soc. 20 No. 219, (1979).

[A.7] S. Arimoto: 情報理論，共立出版 (1976).

[A.8] V. I. Arnold and A. Avez: Ergodic problems of classical mechanics, Math. Phys. Monograph Seri., Benjamin, New York (1968). (吉田訳，古典力学のエルゴード問題，吉岡書店 (1967)).

[A.9] R. Ash: Information Theory, Interscience (1965).

[B.1] A. V. Balakrishnan: Applied Functional Analysis, Spriger-Verlag, (1976).

[B.2] A. T. Bharucha-Reid: Elements of the Theory of Markov Processes and their Applications, McGraw-Hill, (1960).

[B.3] A. T. Bharucha-Reid (Ed): Random Integral Equations, Acad. Press, (1972).

[B.4] A. T. Bharucha-Reid (Ed): Probablistic Methods in Applied Mathematics, Acad. Press, Vol. 1 (1968), Vol. 2 (1970) and Vol. 3 (1973).

[B.5] A. T. Bharucha-Reid (Ed): Probablistic Analysis and Related Topics, Acad. Press, Vol. 1 (1978) and Vol. 2 (1979).

[B.6] A. T. Bharucha-Reid: Approximate Solution of Random Equations, North-Holland, (1979).

[B.7] P. Billingsley: Ergodic Theory and Information, Wiley, New York, (1965). (渡辺・十時訳, 確率論とエントロピー, 吉岡書店 (1968)).

[B.8] L. Breiman: On achieving channel capacity in finite-memory channels, Illinois J. Math. 4 (1960), 246-252.

[C.1] I. Csiszár: Information-type measures of difference of probability distributions and indirect observations, Studia Sci. Math. Hunger., 2 (1967), 299-318.

[D.1] A. Deluca and S. Termini: A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory, Inf. Contr. 20 (1972), 301-312.

[D.2] A. Deluca and S. Termini: Entropy of L-fuzzy sets, Inf. Contr. 24 (1974), 55-73.

[D.3] J. L. Doob: Stochastic Processes, J. Wiley, (1953).

[F.1] A. Feinsten: Foundations of Information Theory, Mcgraw-Hill, New York, (1958).

[F.2] A. Feinstein: On the coding theorem and its converse for finite memory channels, Inf. Contr. 2 (1959), 25-44.

[G.1] I. M. Gelfand and A. M. Yaglom: Calculation of the amount of information about a random function contained in another such function, Amer. Math. Soc. Transl. 12 (1959), 199-246.

[G.2] S. G. Ghurye: Information and Sufficient sub-fields, Ann. Math. Stat. 38 (1968), 2056-2066.

[G.3] R. R. Goldberg: Fourier Transforms, Cambridge Univ. Press, (1961).

[G.4] S. Guiaşu: Information Theory with Applications, Mcgraw-Hill, (1977).

[H.1] P. R. Halmos and L. J. Savage: Application of the Radon-Nikodym theorem to the theory of sufficient statistics, Ann. Math. Statistics 20 (1949), 225-241.

[H.2] P. R. Halmos: Measure Theory, New York, Van Nostrand, (1950).

[H.3] P. R. Halmos; Lectures on Ergodic Theory, Chelsea New York, (1958).

[H.4] P. R. Halmos: Entropy in ergodic theory, Univ. Chicago, (1959).

[H.5] H. Hatori: A note on the entropy of a continuous distribution, Kodai Math. Semi. Rep. 10 (1958), 172-176.

[H.6] Y. Horibe: A note on entropy metrics, Inf. Contr. 22 (1973), 403-404.

[H.7] H. P. Hsu: Fourier Analysis, Simon & Schuster Inc., (1970).

[H.8] K. Huang: Statistical mechanics, J. Wiley, (1963).

[I.1] S. Ikeda: On characterization of the Kullback-Leibler mean nformation for continuous probability distributions, Ann. Inst. Statist. Math. 14 (1962), 73-79.

[I.2] N. Ikeda and S. Watanabe: Stochastic Differential Equations and Diffusion Processes, North-Holland, (1981).

[I.3] S. Ikehara and O. Hirota: 情報理論入門—情型科学の誕生と M. I. T., 啓学出版 (1983).

[I.4] R. S. Ingarden and K. Urbanik: Information without probability, Colloqium Math. 9 (1962), 131-150.

[I.5] R. S. Ingarden: Simplified axioms for information without probability, Prace Matematyczne 9 (1965), 273-282.

[I.6] S. Ito: ルベーグ積分入門, 裳華房 (1963).

[I.7] K. Ito: 確率論, 岩波書店 (1953).

[J.1] E. T. Jaynes: Information theory and statistical mechanics, Phys. Rev. 106 (1957), 620-630 and 108 (1957), 171-182.

[K.1] Y. Kakihara and H. Umegaki: Harmonic analysis and Information channels, 1978-Sem. Appl. Func. Anal. (1978) 19-23.

[K.2] S. Kakutani: On equivalence of Infinite product measures, Ann. Math. 49 (1948), 214-224.

[K.3] G. Kallianpour: On the amount of information contained in a σ-field, in Contribution to Probability and Statistics, Stanford Univ. Press, (1960).

[K.4] T. Kato: 位相解析, 共立出版 (1966).

[K.5] T. Kawata: Fourier analysis ın probability theory, Acad. Press, (1972).

[K.6] J. K. Kelley: General Topology, von Nostrand, (1955). (児玉訳, 位

相空間論，吉岡書店（1969)).

[K.7] J. L. Kelly: Averaging operators on $C_\infty(X)$, Illinois J. Math. 2 (1958), 214-223.

[K.8] A. I. Khinchin: Mathematical Foundations of Information Theory, Dover, New York (English translation), (1958).

[K.9] J. Knopemacher: On measures of fuzziness, J. Math. Anal. Appl. 49 (1975), 529-534.

[K.10] A. N. Kolmogorov: On the representation of continuous functions of many variables by superpositions of continuous functions of one variable and addition, Dokl. Akad. Nauk 114 (1957) 679-681.

[K.11] S. Koshi: 測度と積分，共立出版（1977).

[K.12] K. Kunisawa and H. Umegaki: 情報理論の進歩，岩波書店（1965).

[K.13] K. Kunisawa: エントロピーモデル，日科技連（1975).

[K.14] K. Kunisawa: 情報理論 I，共立出版（1983).

[K.15] S. Kullback and R. A. Leibler: On information and sufficiency, Ann. Math. Statistics 22 (1951), 79-86.

[L.1] M. Loève: Probability theory, van Nostrand, (1960).

[M.1] S. Maeda, 函数解析，森北出版（1974).

[M.2] N. F. G. Martin and J. W. England: Mathematical Theory of Entropy, Addison-Wesley, (1981).

[M.3] B. McMillan; The basic theorem of information theory, Ann. Math. Stat. 24 (1953), 196-219.

[N.1] S. Nabeya: 数理統計学，共立出版（1979).

[N.2] M. Nakamura and H. Umegaki: A remark of theorems of Stone and Bochner, Proc. Japan Acad. 27 (1951), 506-507.

[N.3] M. Nakamura and H. Umegaki: Heisenberg's commutation reration and the Plancherel theorem, Proc. Japan. Acad. 37 (1961), 239-242.

[N.4] Y. Nakamura; Measure theoritic construction for information theory, Kodai Math. Sem. Rep. 21 (1969), 133-150.

[N.5] Y. Nakamura: Ergodicity and capacity of information channels with noise sources, J. Math. Soc. Japan 27 (1975), 213-221.

[N.6] Y. Nakamura: On an extension of K-L expansion and an application to classification problem, 1978-Sem. Appl. Funct. Anal. (1978) 1-7.

[N.7] O. B. Nielsen: Subfields and loss of information, Wahrschernlichkeits theories 2 (1964), 369-379.

[O.1] W. Ochs: Basic properties of the generalized Boltzmann-Gibbs-Shannon entrrpy, Rep. Math. Phys. 9 (1976), 135-155 and 331-354.
[O.2] M. Ohya: 情報とゆらぎ, 数理科学 188 (1979), 37-42.
[O.3] M. Ohya: Quantum ergodic channels in operator algebras, J. Math. Anal. Appl. 84 (1981), 318-328.
[O.4] M. Ohya: On compound state and mutual information in quantum information theory, IEEE Inf. Theor. 29 (1983), 770-774.
[O.5] D. S. Ornstein: Bernoulli shifts with the same entropy are isomorphic, Adv. in Math. 4 (1970), 337-352.
[O.6] D. S. Ornstein and P. Shields: An uncountable family of K-automorphisms, Adv. in Math. 10 (1973), 63-88.
[O.7] J. C. Oxtoby: Ergodic sets, Bull. Amer. Math. Soc. 58 (1952), 116-136.

[P.1] W. Parry: Topics in ergodic theory, Cambridge Univ. Press, (1981).
[P.2] K. R. Parthasarathy: On the integral representation of the rate of transmission of a stationary channel, Ill. J. Math. 5 (1961), 299-305.
[P.3] K. R. Parthasarathy: Probability measures on metric spaces, Academic Press, (1967).
[P.4] R. T. Prosser: The ε-entropy and ε-capacity of certain time-varying channels, J. Math. Anal. Appl. 16 (1966), 553-573.
[P.5] R. T. Prosser and W. L. Root: Determinable classes of channels, J. Math. and Mech. 16 (1966), 365-398.

[R.1] M. Reed and B. Simon: Methods of Modern Mathematical Physics Vol. I: Functional Analysis, Academic Press, (1973).
[R.2] M. Reed and B. Simon: Methods of Modern Mathematical Physics Vol. II: Fourier Analysis, Self-Adjointness, Academic Press, (1975).
[R.3] A. Rényi: Probability Theory, North-Holland, Amsterdam, (1970).

[S.1] H. Sato: 情報理論, 裳華房 (1973).
[S.2] R. Schatten: A theory of cross-spaces. Ann. Math. Studies No. 26,

princeton Univ. press 1950.
[S.3] R. Schatten: Norm ideals of completely continuous operators, Springer-verlag, (1970).
[S.4] C. R. Shannon: A mathematical theory of communication, Bull System Tech. J. 27 (1948), 379-423 and 623-656.
[S.5] P. Shields: The theory of Bernoulli Shifts, Univ. of Chicago Press, (1973).
[S.6] Ja. G. Sinai: On the concept of entropy for dynamical systems, Doki, Akad. Nauk. SSSR, 124 (1959), 768-771, (Russian).
[S.7] M. Sugeno: Fuzzy measures and fuzzy integrals: A surbey, in Fuzzy Automata and Decision Process (ed by M. M. Gupta et al.), North-Holland, Amsterdam, (1977).

[T.1] K. Takano: On the basic Theorems of information theory, Ann. Instit. Statist. Math. Tokyo 9 (1958), 53-77.
[T.2] Z. Takeda: 代数系と符号理論, 槇書店 (1977).
[T.3] Y. Taki: 情報論 I, 岩波全書 (1978).
[T.4] T. Totoki: エルゴード理論入門, 共立出版 (1971).
[T.5] H. Tverberg: A new derivation of the information function, Math. Scand 6 (1958), 297-298.

[U.1] Y. Uesaka: 情報数学の基礎, 培風館 (1977).
[U.2] Y. Uesaka, T. Funada and N. Kimata: 情報科学の基礎, 培風館 (1979).
[U.3] H. Umegaki: Positive definite function and direct product Hilbert space, Tohoku. M. J. 7 (1955), 206-211.
[U.4] H. Umegaki: A functional method for stationary channels, Kodai Math. Sem. Rep. 16 (1964), 27-39.
[U.5] H. Umegaki: Operator methods for abstract harmonic analysis, MRC-Tech. Rep. Univ. Wisc. 6, (1967).
[U.6] H. Umegaki: Representation and extremal properties of averaging operators and their applications to information channels, J. Math. Anal. Appl. 25 (1969), 41-73.
[U.7] H. Umegaki: Absolute continuity of information channels, J. Multiva. Anal. 4 (1974), 382-400.
[U.8] H. Umegaki and A. T. Bharucha-Reid: Banach Space valued random variable and tensor product banach spaces, J. Math. Ann. Appl. 31

(1970), 49-67.

[U.9] H. Umegaki: General treatment of alphabet message space and integral representation of entropy, Kodai Math. Sem. Rep. 16 (1964), 18-26.

[U.10] H. Umegaki and Y. Nakamura: 情報理論とエントロピー, 数理科学, 5 (1970).

[U.11] H. Umegaki: 情報路と作用素, 数理科学 6 (1971).

[U.12] H. Umegaki: 情報理論と作用素, 数理科学 110 (1972), 29-35.

[U.13] H. Umegaki, M. Ohya and F. Hiai: エントロピーと情報, 日本物理学会誌 35 (1980), 241-252.

[U.14] H. Umegaki: A functional method on amount of entropy, Kodai Math. Sem. Rep., 15 (1963), 162-175.

[V.1] J. von Neumann: Collected Works,
Vol. 1: Logic, Theory of Sets and Quantum Mechanics,
Vol. 2: Operators, Ergodic theory and almost periodic functions in a group,
Vol. 3: Rings of operators,
Vol. 4: Continuous geometry and other topics,
Vol. 5: Design of computer, theory of automata and numerical analysis,
Vol. 6: Theory of Games, Astrophysics, Hydrodynamics and Meteorology.

[W.1] P. Walter: Ergodic theory—introductory lectures, Lecture notes in Math. No. 458, Springer, Berlin, (1975).

[W.2] S. Watanabe: Knowing and Guessing, Wiley, New York, (1969).

[W.3] N. Wiener: Cybernetics, 2 nd, MIT Press, (1961).

[Y.1] Y. Yamazaki: 無限次限空間の測度 (上, 下), 紀伊国屋書店 (1978).

[Y.2] K. Yosida: Functional Analysis, 2 nd ed. Springer-Verlag, (1968).

[Y.3] J. F. Young: Information Theory, Butterworth, (1971).

[Z.1] L. A. Zadeh: Fuzzy sets, Inf. Contr. 8 (1965), 338 353

[Z.2] L. A. Zadeh: Probability measures of fuzzy events, J. Math. Anal. Appl. 23 (1968), 421-427.

索　引

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

R

S

T

U

V

W

Y

Z

Memorandum

Memorandum

著者紹介

梅垣(うめがき) 壽春(ひさはる)

昭和 24 年　東北帝国大学理学部数学科卒業
専　　攻　応用函数解析，エントロピー論，信号解析，非可換解析
現　　在　東京工業大学理学部教授　理学博士

大矢(おおや) 雅則(まさのり)

昭和 45 年　東京大学理学部物理学科卒業
昭和 51 年　Rochester 大学大学院修了
専　　攻　エントロピー論，数理物理学，函数解析
現　　在　東京理科大学理工学部助教授　Ph. D（理学博士）

情報科学講座 A・2・6
確率論的エントロピー
—情報理論の函数解析的基礎 1—

検印廃止

定価 2700 円

昭和 58 年 12 月 20 日　初版 1 刷発行
昭和 59 年 4 月 15 日　初版 2 刷発行

編　集　北川敏男
代表者　喜安善市
発行者　南條正男
東京都文京区小日向 4 丁目 6 番 19 号
印刷者　大久保絢史
東京都新宿区市ヶ谷本村町 27 番地

NDC 007.1, 415.5, 548.1

発行所　東京都文京区小日向 4 丁目 6 番19号
電話 東京 947 局 2511 番 （代表）
郵便番号 113 振替 東京 1-57035番
共立出版株式会社

印刷・新日本印刷　製本・関山製本　Printed in Japan

社団法人
自然科学書協会
会　員

ISBN 4-320-02209-2